



Dirichlet is Natural

Vincent Danos, Ilias Garnier

► To cite this version:

Vincent Danos, Ilias Garnier. Dirichlet is Natural. MFPS 31 - Mathematical Foundations of Programming Semantics XXXI, Jun 2015, Nijmegen, Netherlands. pp.137-164, 10.1016/j.entcs.2015.12.010 . hal-01256903

HAL Id: hal-01256903

<https://hal.science/hal-01256903>

Submitted on 15 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Dirichlet is natural

Vincent Danos¹

*Département d'Informatique
École normale supérieure
Paris, France*

Ilias Garnier²

*School of Informatics
University of Edinburgh
Edinburgh, United Kingdom*

Abstract

Giry and Lawvere's categorical treatment of probabilities, based on the probabilistic monad G , offer an elegant and hitherto unexploited treatment of higher-order probabilities. The goal of this paper is to follow this formulation to reconstruct a family of higher-order probabilities known as the *Dirichlet process*. This family is widely used in non-parametric Bayesian learning.

Given a Polish space X , we build a family of higher-order probabilities in $G(G(X))$ indexed by $M^*(X)$ the set of non-zero finite measures over X . The construction relies on two ingredients. First, we develop a method to map a zero-dimensional Polish space X to a projective system of finite approximations, the limit of which is a zero-dimensional compactification of X . Second, we use a functorial version of Bochner's probability extension theorem adapted to Polish spaces, where consistent systems of probabilities over a projective system give rise to an actual probability on the limit. These ingredients are combined with known combinatorial properties of Dirichlet processes on finite spaces to obtain the Dirichlet family \mathcal{D}_X on X . We prove that the family \mathcal{D}_X is a *natural transformation* from the monad M^* to $G \circ G$ over Polish spaces, which in particular is continuous in its parameters. This is an improvement on extant constructions of \mathcal{D}_X [17,26].

Keywords: probability, topology, category theory, monads

1 Introduction

It has been argued that exact bisimulations between Markovian systems are better conceptualized using the more general notion of bisimulation metrics [29]. This is because there are frequent situations where one can only estimate the transition probabilities of a Markov chain (MC).³ Such uncertainties lead one naturally to using a metric-based notion of approximate equivalence as a more robust way of comparing processes than exact bisimulations. Here, we wish to take a new look at

¹ vincent.danos@ens.fr

² igarnier@inf.ed.ac.uk

³ Even though the existence of symmetries in physical systems can sometimes lead to exact bisimulations which depend only on structure and not on the actual values of transition probabilities [28]. There are attempts, parallel to bisimulation metrics, at defining robustly the satisfaction of a temporal logic formula [14]

this issue of uncertainty in the model and suggest a novel and richer framework to deal with it. We keep the idea of using a robust means of comparison (typically the Kantorovich or Prohorov metrics lifted to MCs), but we add a second idea: namely to introduce a way of quantifying the uncertainty in the chains being compared.

To quantify uncertainty in the Markov chains, we propose to explore in the longer term concepts of “uncertain Markov chains” as elements of type $X \rightarrow G^2(X)$, where X is an object of **Pol**, the category of Polish spaces (separable and completely metrisable spaces) and G is the Giry probability functor. This is to say that the chain takes values in “random probabilities” (ie probabilities of probabilities).⁴

This natural treatment of behavioural uncertainty in probabilistic models will allow one to formulate a notion of (Bayesian) learning and therefore to obtain notions of 1) models which can learn under observations and 2) of behavioural comparisons which can incorporate data and reduce uncertainty. Bisimulation metrics between processes become random variables and learning should decrease their variability.

One needs to set up a sufficiently general framework for learning under observation within the coalgebraic approach. Learning a probability in a Bayesian framework is naturally described as a (stochastic) process of type $G^2(X) \rightarrow G^2(X)$ (so $G^2(X) \rightarrow G^3(X)$ really!) driven by observations. For finite X s this setup poses no difficulty, but for more general spaces, one needs to construct a computational handle on $G^2(X)$ - the space of uncertain or higher-order probabilities. This is what we do in this paper.

To this effect, we build a theory of Dirichlet-like processes in **Pol**. Dirichlet processes [1,16] form a family of elements in $G^2(X)$ indexed by finite measures over X [1, p.17]⁵ and which is closed under Bayesian learning.

Integral to our construction is a method of “decomposition/recomposition” which allows us to build higher-probabilities via finite approximations of the underlying space (the limit of which lead to a compactification of the original space). In order to lift finite higher-probabilities we use a bespoke extension theorem of the Kolmogorov-Bochner type in **Pol** (Sec. 2.3). Kolmogorov consistent assignments of probabilities on finite partitions of measurable spaces (or finite joint distributions of stochastic processes) can be seen systematically as points in the image under G of projective (countable co-directed) diagrams in **Pol**.

Using the above we show that Dirichlet-like processes in **Pol** can be seen as natural transformations from M^* (the monad of non-zero finite measures on **Pol**) to G^2 built up from finite discrete spaces. The finite version of naturality goes under the name of “aggregation laws” in the statistical literature and can be traced back to the “infinite divisibility” of the one building block, namely the Γ distribution. (This opens up the possibility of an axiomatic version of the construction presented here, see conclusion.)

⁴ Another possibility is to consider uncertain chains as elements of $G(X \rightarrow G(X))$, but, unless X is compact, this takes us outside of **Pol**.

⁵ Eg as for Poisson point processes.

2 Notations & basic facts

We provide a primer of general topology as used in the paper in Appendix A. A useful reference on the matter is [11]. Weak convergence of probability measures is treated in [7, 27].

2.1 Finite measures on Polish spaces and the Giry monad

Weak topology

A measure P on a topological space X is a positive countably additive set function defined on the Borel σ -algebra $\mathcal{B}(X)$ verifying $P(\emptyset) = 0$. We will only consider finite measures on Polish spaces, i.e. $P(X) < \infty$. When $P(X) = 1$, P is a *probability measure*. We write $G(X)$ for the space of all probability measures over X with the *weak topology* [7, 27], the initial topology for the family of evaluation maps $EV_f = P \mapsto \int_X f dP$ where f ranges in $C_b(X)$ and where $(C_b(X), \|\cdot\|_\infty)$ is the Banach space of real-valued continuous bounded functions over X with the sup norm. A neighbourhood base for a measure $P \in G(X)$ is given by the sets

$$N_P(f_1, \dots, f_n, \epsilon_1, \dots, \epsilon_n) = \left\{ Q \left| \left| \int f_i dP - \int f_i dQ \right| < \epsilon_i, 1 \leq i \leq n \right. \right\}$$

where $f_i \in C_b(X)$, $\epsilon_i > 0$. One can restrict w.l.o.g. to the subset of real-valued bounded uniformly continuous functions, noted $U_b(X)$. Importantly, if X is Polish the weak topology on $G(X)$ is also Polish (see e.g. Parthasarathy, [27] Chap. 2.6) and metrisable by the Wasserstein-Monge-Kantorovich distance [31]. We denote the convergence of a sequence $(P_n \in G(X))_{n \in \mathbb{N}}$ to $P \in G(X)$ in the weak sense by $P_n \rightharpoonup P$. The “Portmanteau” theorem ([7], Theorem 2.1) asserts that $P_n \rightharpoonup P$ is equivalent to $P_n(B) \rightarrow P(B)$ for all P -continuity sets B , i.e. Borel sets s.t. $P(\partial B) = 0$. P -continuity sets form a Boolean algebra ([27], Lemma 6.4). The *support* of a probability $P \in G(X)$ is noted $\text{supp}(P)$ and is defined as the smallest closed set such that $P(\text{supp}(P)) = 1$. For X, Y Polish and $P \in G(X), Q \in G(Y)$, we write $P \otimes Q \in G(X \times Y)$ the *product probability*, so that $(P \otimes Q)(B_X \times B_Y) = P(B_X)Q(B_Y)$.

Giry monad

The operation G can be extended to a functor $G : \mathbf{Pol} \rightarrow \mathbf{Pol}$ compatible with the Giry monad structure (G, δ, μ) [19]. For any continuous map $f : X \rightarrow Y$ we set $G(f)(P) = B \in \mathcal{B}(Y) \mapsto P(f^{-1}(B))$, i.e. $G(f)(P)$ is the pushforward measure. For a given X , $\delta_X : X \rightarrow G(X)$ is the Dirac delta at x while $\mu_X : G^2(X) \rightarrow G(X)$ is defined as averaging: $\mu_X(P) = B \in \mathcal{B}(X) \mapsto \int_{G(X)} EV_B dP$ where $EV_B = Q \in G(X) \mapsto Q(B)$ evaluates a probability on the Borel set B . We have the “change of variables” formula: for all $P \in G(X)$, $f : X \rightarrow Y$ and $g : Y \rightarrow \mathbb{R}$ bounded measurable, $\int_Y g dG(f)(P) = \int_X g \circ f dP$. Finally, G preserves surjectivity, injectivity and openness:

Lemma 2.1 (i) $f : X \rightarrow Y$ is injective if and only if $G(f)$ is injective;
(ii) f is surjective if and only if $G(f)$ is surjective.

(iii) If f is an embedding, so is $G(f)$.

Proof. We recall that elements of $G(X)$ for X Polish verify the *Radon property*: for all Borel set $B \in \mathcal{B}(X)$ and all $P \in G(X)$, $P(B) = \sup \{P(K) \mid K \subseteq B, K \text{ compact}\}$ (see [9], Chap. 7). (i) Let f be an injective continuous map. Let $P, Q \in G(X)$ be such that $P(B) \neq Q(B)$ for some Borel set B . Then, there must exist a compact $K \subseteq B$ such that $P(K) \neq Q(K)$. The set $f(K)$ is compact, hence Borel; by injectivity $G(f)(P)(f(K)) = P(K)$ and similarly for Q , therefore $G(f)(P)(f(K)) \neq G(f)(Q)(f(K))$. Conversely, if $G(f)$ is injective then it is in particular injective on the set $\{\delta_x \mid x \in X\} \subseteq G(X)$, therefore f is injective. (ii) Let f be surjective continuous. Let $Q \in G(Y)$ be some probability. By the measurable selection theorem [32], there exists a measurable function $g : Y \rightarrow X$ such that $g(y) \in f^{-1}(y)$, which implies $f \circ g = \text{id}_Y$. Let P be the pushforward measure of Q through g , i.e. $P(B) \triangleq Q \circ g^{-1}$. By surjectivity of f , $P(X) = 1$, therefore $P \in G(X)$. The identity $f \circ g = \text{id}_Y$ entails $G(f)(P) = Q$. Conversely, assume $G(f)$ is surjective. Since $\{\delta_y \mid y \in Y\} \subseteq G(Y)$, there must exist for each y a $P_y \in G(X)$ such that $\delta_y(y) = (P_y \circ f^{-1})(y) > 0$, therefore f is surjective. (iii) Assume f is an embedding. Let $N_P(g_1, \dots, g_n, \epsilon_1, \dots, \epsilon_n)$ be some basic neighbourhood of some $P \in G(X)$, and let $P' \in N_P$ be in the neighbourhood of P , i.e. $|\int_X g_i dP - \int_X g_i dP'| < \epsilon_i$ for $1 \leq i \leq n$. Note that since f is an embedding, for each $g_i \in C_b(X)$ there exists a $g'_i \in C_b(f(X))$ verifying $g'_i(f(x)) = g_i(x)$. Therefore:

$$\begin{aligned} \left| \int_Y g'_i dG(f)(P) - \int_Y g'_i dG(f)(P') \right| &= \left| \int_X g'_i \circ f dP - \int_X g'_i \circ f dP' \right| \\ &= \left| \int_X g_i dP - \int_X g_i dP' \right| < \epsilon_i \end{aligned}$$

□

Finite measures

The set of all finite non-negative Borel measures on a Polish space, noted $M(X)$, is a Polish space when endowed with the weak topology ([9] Theorem 8.9.4). $M : \mathbf{Pol} \rightarrow \mathbf{Pol}$ is a functor extending G , mapping continuous functions to the corresponding pushforward morphism. The monad multiplication μ_X can be conservatively extended to a morphism from $M^2(X)$ to $M(X)$ by defining $\mu_X(P) = B \in \mathcal{B}(X) \mapsto \int_{M(X)} EV_B dP$. The everywhere zero measure, noted $\mathbf{0}$, is an element of $M(X)$ that we might want to exclude: $M(X)$ being Hausdorff implies that the set of nonzero measures $M^*(X) \triangleq M(X) \setminus \{\mathbf{0}\}$ is open, hence G_δ , hence Polish as a subspace of $M(X)$. A measure $Q \in M(X)$ is *strictly positive* if for all nonempty open sets $U \subseteq X$, $Q(U) > 0$. Equivalently, Q is strictly positive if and only if $\text{supp}(Q) = X$.

Lemma 2.2 *Strictly positive finite measures on a Polish space X form (when they exist) a Polish subspace of $M(X)$. We denote this subspace by $M^+(X)$.*

Proof. It is sufficient to show that $M^+(X)$ is a G_δ set in $M(X)$. Let $\{O_n\}_{n \in \mathbb{N}}$ be a countable base of X . Strict positivity of a measure Q is equivalent to having $Q(O_n) > 0$ for all nonempty O_n , therefore $M^+(X) = \bigcap_n \{Q \in M(X) \mid Q(O_n) > 0\}$.

Clearly $\{Q \mid Q(O_n) = 0\}$ is closed in the weak topology, therefore $M^+(X)$ is a G_δ , and forms a Polish subspace of $M(X)$. \square

Summing up, we have for X Polish the following inclusions of Polish spaces of finite measures:

$$M^+(X) \subseteq M^*(X) \subseteq M(X)$$

Note also that M and M^* are endofunctors on **Pol** but M^+ is *not*, unless one restricts to the subcategory of epimorphisms.

Normalisation of measures

We note $\nu_X : M^*(X) \rightarrow G(X)$ the continuous map taking any measure $Q \in M^*(X)$ to its normalisation $\nu_X(Q) \triangleq B \in \mathcal{B}(X) \mapsto Q(B)/|Q|$, where $|Q| \triangleq Q(X)$ is the total mass of the measure. ν_X verifies an useful property:

Lemma 2.3 $\nu : M^* \Rightarrow G$ is natural.

Proof. Let $f : X \rightarrow Y$ be a continuous map. We have:

$$(G(f) \circ \nu_X)(Q) = \nu_X(Q) \circ f^{-1} = \frac{Q \circ f^{-1}}{Q(X)} = \frac{Q \circ f^{-1}}{Q(f^{-1}(Y))} = \nu_Y \circ M^*(f)(Q)$$

\square

Densities and convolution

The Radon-Nikodym theorem asserts that measures in $G(\mathbb{R})$ absolutely continuous with respect to the Lebesgue measure admit integral representations such that $P(A) = \int_A f dx$. In this case P is said to have *density* f with respect to the Lebesgue measure. f is sometimes noted $\frac{dP}{d\lambda}$, where λ denotes Lebesgue. For $P, Q \in G(\mathbb{R})$ with respective densities w.r.t. Lebesgue f_P, f_Q , the convolution product of P and Q is defined to be the measure $P * Q$ having density $f_{P*Q}(x) = \int_{\mathbb{R}} f_P(x) f_Q(x - t) dt$ (see Kallenberg [22], Lemma 1.28).

Finitely supported measures

When X is a finite, discrete space such that $X = \{x_1, \dots, x_n\}$, $G(X)$ is in bijection with the simplex $\Delta_n \subseteq \mathbb{R}^n$, where $\Delta_n = \{(p_1, \dots, p_n) \in \mathbb{R}^n \mid p_i \geq 0, \sum p_i = 1\}$. Notice that Δ_n is an $n - 1$ dimensional space. $M(X)$ corresponds to the positive orthant, noted $\mathbb{R}_{\geq 0}^n$. Since for X finite $G(X)$ is (topologically) a subspace of a finite dimensional vector space, it is homeomorphic to $\Delta_n \cap \mathbb{R}^n$ while the topology of $M(X)$ corresponds to that of $\mathbb{R}_{\geq 0}^n \cap \mathbb{R}^n$. If we note \mathfrak{n} the n -element set, we in particular have the trivial identities $M(\mathfrak{n}) = \mathbb{R}_{\geq 0}^n$ and $M(\mathfrak{m}) \times M(\mathfrak{n}) = M(\mathfrak{m} + \mathfrak{n})$.

2.2 Projective limits of topological spaces

Many of our theorems will deal with spaces obtained as *projective limits* (also known as *inverse limits* or *cofiltered limits*) of topological spaces. These topological projective limits are defined as adequate topologisations of projective limits in **Set**, the usual category of sets and functions.

Let (I, \leq) be a directed partially ordered set seen as a category and let $D : I^{op} \rightarrow \mathbf{Set}$ be a cofiltered \mathbf{Set} diagram. The projective limit of D is a terminal cone $(\lim D, \pi_i)$ over D where $\lim D$ is the set

$$\lim D \triangleq \{x \mid D(i \leq j)(\pi_j(x)) = \pi_i(x)\} \subseteq \prod_i D(i)$$

and the $\pi_i : \prod_j D(j) \rightarrow D(i)$ are the canonical projections. Notice that D is contravariant from I to \mathbf{Set} . As emphasised in the definition, $\lim D$ is the subset of the cartesian product $\prod_i D(i)$ containing all sequences of elements that respect the constraints imposed by the diagram D . The elements of $\lim D$ are called *threads* and the maps $D(i \leq j) : D(j) \rightarrow D(i)$ are the *bonding maps*. Of course, $\lim D$ can be empty (see [34] for a short example). A sufficient condition to ensure non-emptiness of the limit is to consider functors D where I is countable and the bonding maps are surjective [6]. As a convenience, we will note those bonding maps as $\pi_{ij} \triangleq D(i \leq j)$, and we write countable cofiltered surjective diagrams *ccd* for short.

Writing $U : \mathbf{Top} \rightarrow \mathbf{Set}$ for the underlying set functor, cofiltered limits in \mathbf{Top} for diagrams $D : I^{op} \rightarrow \mathbf{Top}$ are obtained by endowing the \mathbf{Set} limit of $U \circ D$ with the initial topology for the canonical projections $\{\pi_i\}_{i \in I}$. The following useful additional fact follows by considering $\lim D$ as the intersection of the (closed) subsets of $\prod_i D(i)$ satisfying $D(i \leq j)(\pi_j(x)) = \pi_i(x)$ for all pairs (i, j) s.t. $i \leq j$.

Lemma 2.4 ([11], Ch. 1, §8.2, Corollaire 2) *$\lim D$ is a closed subset of $\prod_i D(i)$.*

2.3 The Bochner extension theorem

The construction of a stochastic process given a system of consistent finite-dimensional marginals is an important tool in probability theory, a classical example being the construction of the Brownian motion using the *Kolmogorov extension theorem* [25]. Besides Kolmogorov's there are many other variants, collectively called *Bochner extension theorem* [24]. They differ in the amount of structure of the space over which probabilities are considered (measurable, topological or vector spaces) – and we will make crucial use of the Bochner extension theorem for Polish spaces, which admits a particularly elegant presentation.

Theorem 2.5 *For all D a ccd in \mathbf{Pol} , $G(\lim D) \cong \lim G \circ D$. We denote by $\text{bcn} : \lim G \circ D \rightarrow G(\lim D)$ this homeomorphism.*

In words, the Bochner extension theorem states that any projective family of probabilities that satisfy the diagram constraints (elements of $\lim G \circ D$) can be uniquely lifted to a probability over the limit space (elements of $G(\lim D)$) – and what's more, this extension is a homeomorphism! This presentation of the Bochner extension seems not to be well-known: a similar statement is given in Metivier ([24], Theorem 5.5) in the case of locally compact spaces, which intersects but does not include Polish spaces; Fedorchuk proves the continuity of G on the class of compact Hausdorff spaces in [15] while more recently Banach [4] provides an extension theorem in the more general setting of Tychonoff spaces, using properties of the Stone-Čech compactification.

3 Zero-dimensional Polish spaces and their properties

It is natural in applications to consider finitary approximations of stochastic processes. Accordingly, the correctness of such approximations should correspond to some kind of limiting argument, stating that increasingly finer approximations yield in some suitable sense the original object. In view of the Bochner extension theorem, it suffices to consider as input a projective family of probabilities supported by the finitary approximants of the underlying space. However, the very same theorem tells us that we can only obtain by this means probabilities on a projective limit of finite spaces (also called *profinite* spaces), a rather restrictive class:

Proposition 3.1 *A space is a countable projective limit of finite discrete spaces if and only if it is a compact, zero-dimensional Polish space.*

The proof can be found under a slightly different terminology in Borceux & Janelidze [10], where it is shown that these spaces correspond to *Stone spaces* – indeed, profinite spaces are exactly the spaces homeomorphic to the Stone dual of their Boolean algebra of clopen sets! As the proof of this proposition is quite enlightening for the developments to come, we provide it here.

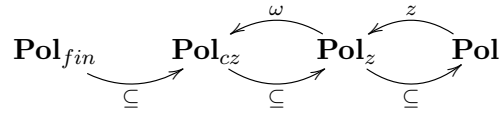
Proof. Let $D : I^{op} \rightarrow \mathbf{Pol}_{fin}$ be a *ccd* of finite spaces. Polishness of $\lim D$ comes from the closure of \mathbf{Pol} under countable limits. Finite spaces are compact and by Tychonoff’s theorem so is $\prod_i D(i)$. Lemma 2.4 asserts that $\lim D$ is closed in this compact product, hence $\lim D$ is itself compact. Recall that $\lim D$ has the initial topology for the canonical projections maps $\pi_i : \lim D \rightarrow D(i)$, therefore a base of $\lim D$ is constituted of finite intersections of prebase opens $\pi_i^{-1}(X_i)$, for $X_i \subseteq D(i)$. Since the $D(i)$ are discrete, any of their subsets is clopen and so are the prebase opens; we conclude by noticing that a finite intersection of clopen sets is again clopen.

Conversely, let Z be a compact zero-dimensional Polish space. As Z is zero-dimensional Polish, its topology is generated by a countable base of clopen sets. Since Z is compact, each clopen can be written as a finite union of base clopens. Therefore its Boolean algebra of clopens $Clo(Z)$ is also generated by the same countable base, and is itself countable. Note that $Clo(Z)$ does not depend on the choice of the base! Let us consider the set $\mathcal{I}(Z)$ of all finite clopen partitions of Z . For any $i \in \mathcal{I}(Z)$, there exist by assumption a continuous surjective quotient map $f_i : Z \rightarrow i$. Since i is discrete, the fibres of f_i are clopen. Note that $\mathcal{I}(Z)$ is also countable. $\mathcal{I}(Z)$ is partially ordered by partition refinement: for all $i, j \in \mathcal{I}(Z)$, we write $i \leq j$ if there exists a surjective “bonding” map $f_{ji} : j \rightarrow i$ such that $f_{ji} \circ f_j = f_i$ (any such map, if it exists, is unique). $\mathcal{I}(Z)$ is also directed by considering pairwise intersections of the cells of any two partitions. The system of finite discrete quotients of Z together with the bonding maps f_{ji} clearly defines a *ccd* that we write $D : \mathcal{I}(Z)^{op} \rightarrow \mathbf{Pol}_{fin}$, mapping each element of $\mathcal{I}(Z)$ to itself and the partial order of refinement to the bonding maps. Therefore, there exists a limit cone $(\lim D, \pi_i)$. By universality of this cone, there exists a unique continuous map $\eta : Z \rightarrow \lim D$ s.t. $f_i = \pi_i \circ \eta$. Let us show that η is an homeomorphism. As $\lim D$ and Z are both compact, it is enough to show that η is a bijection. Recall that $Clo(Z)$ separates points (it contains a base for a Hausdorff topology) therefore

for any $x \neq y \in Z$ we can exhibit two clopen cells separating them, implying that η is injective. Surjectivity of η is a consequence of that of the quotient and bonding maps. \square

We denote by \mathbf{Pol}_{cz} the full subcategory of \mathbf{Pol} where objects are compact and zero-dimensional – by the previous proposition, these spaces are exactly the profinite Polish spaces. From the data of a projective system of finitely supported probabilities, Prop. 3.1 together with Bochner’s extension theorem (Thm. 2.5) only allow us to obtain probabilities supported by such profinite spaces. Our extension of Dirichlet as a natural transformation from the setting of finite spaces to that of arbitrary Polish spaces must therefore imperatively bridge the gap from profinite spaces to arbitrary Polish spaces.

The solution we propose is mediated by zero-dimensional Polish spaces in a decisive way. More precisely, our construction can be framed as the iterative reduction of the extension problem to increasingly smaller subcategories of \mathbf{Pol} (depicted below): the (full) subcategory of zero-dimensional spaces \mathbf{Pol}_z , that of compact zero-dimensional spaces \mathbf{Pol}_{cz} and finally the subcategory \mathbf{Pol}_{fin} of finite Polish spaces. The categorical setting is informally sketched in the following figure:



The two essential operations, highlighted in the figure above, are:

- the *zero-dimensionalisation* Z , which yields a zero-dimensional refinement of a Polish space for which a countable base of the topology has been chosen, and
- the *zero-dimensional Wallman compactification* ω , which yields a compact zero-dimensional Polish space from a zero-dimensional one along, again, a choice of a base of clopens sets.

We attract the attention of the reader on the fact that these operations are a priori *not* functorial. However, as we shall see in the rest of this section, these operations exhibit powerful properties which are sufficient to proceed to the extension.

3.1 Zero-dimensionalisation

Zero-dimensionalisation takes as input a Polish space X along a choice of some countable base \mathcal{F} for X . It produces a Polish zero-dimensional topology on the same underlying set as X , that we denote by $z_{\mathcal{F}}(X)$.

Proposition 3.2 *Let (X, \mathcal{T}_X) be Polish and let \mathcal{F} be a countable base for X . Let $\text{Boole}(\mathcal{F})$ be the Boolean algebra generated by \mathcal{F} . Let $z_{\mathcal{F}}(X)$ be the space which admits $\text{Boole}(\mathcal{F})$ as a base of its topology. $z_{\mathcal{F}}(X)$ verifies the following properties:*

- (i) $z_{\mathcal{F}}(X)$ is Polish;
- (ii) $z_{\mathcal{F}}(X)$ is zero-dimensional;
- (iii) the Borel sets are preserved: $\mathcal{B}(X) = \mathcal{B}(z_{\mathcal{F}}(X))$;

(iv) the identity function $id_{\mathcal{F}} : z_{\mathcal{F}}(X) \rightarrow X$ is continuous.

In order to prove Prop. 3.2 we need some classical facts from descriptive set theory, taken verbatim from Kechris [23], Sec. 13:

Lemma 3.3 *For any Polish space (X, \mathcal{T}_X) and any closed set A , there exists a Polish topology \mathcal{T}_{X_A} so that $\mathcal{T}_X \subseteq \mathcal{T}_{X_A}$, A is clopen in \mathcal{T}_{X_A} and $\mathcal{B}(\mathcal{T}_X) = \mathcal{B}(\mathcal{T}_{X_A})$. Moreover, $\mathcal{T}_X \cup \{O \cap A \mid O \in \mathcal{T}_X\}$ is a base of \mathcal{T}_{X_A} .*

Lemma 3.4 *Let (X, \mathcal{T}_X) be Polish and let $\{\mathcal{T}_{X_n}\}_{n \in \mathbb{N}}$ be a family of Polish topologies on X , then the topology \mathcal{T}_{X_∞} generated by $\cup_n \mathcal{T}_{X_n}$ is Polish. Moreover if $\forall n, \mathcal{T}_{X_n} \subseteq \mathcal{B}(\mathcal{T}_X)$, then $\mathcal{B}(\mathcal{T}_{X_\infty}) = \mathcal{B}(\mathcal{T}_X)$.*

Proof. (Proposition 3.2) For each $O_n \in \mathcal{F}$, let us denote $A_n = X \setminus O_n$. Consider the family of Polish topologies $\{\mathcal{T}_{X_{A_n}}\}_{n \in \mathbb{N}}$, as obtained using Lemma 3.3. Lemma 3.4 entails that the topology generated by $\cup_n \mathcal{T}_{X_{A_n}}$ is Polish. Recall that each $\mathcal{T}_{X_{A_n}}$ has base $\mathcal{T}_X \cup \{O \cap A_n \mid O \in \mathcal{T}_X\}$. Closing $\cup_n \mathcal{T}_{X_{A_n}}$ under finite intersections yields that the topology generated by $\cup_n \mathcal{T}_{X_{A_n}}$ has base $\mathcal{T}_X \cup \{O \cap C \mid O \in \mathcal{T}_X, C \in \text{Boole}(\mathcal{F})\}$. Since \mathcal{F} is a base of \mathcal{T}_X and $\mathcal{F} \subseteq \text{Boole}(\mathcal{F})$, an equivalent base of the topology generated by $\cup_n \mathcal{T}_{X_{A_n}}$ is $\text{Boole}(\mathcal{F})$. By definition, we deduce that the topology of $z_{\mathcal{F}}(X)$ is generated by $\cup_n \mathcal{T}_{X_{A_n}}$.

- (i) Lemma 3.4 entails that the resulting space is indeed Polish. An equivalent base to $\mathcal{T}_X \cup \mathcal{T}_X|_{\mathcal{F}_\delta^c}$ is $\mathcal{F} \cup \mathcal{F}|_{\mathcal{F}_\delta^c}$ and the elements of this base are clopen, hence the resulting space is also zero-dimensional.
- (ii) Zero-dimensionality is a trivial consequence of taking a Boolean algebra as a base.
- (iii) Preservation of Borel sets is a further consequence of Lemma 3.4.
- (iv) Continuity of the identity is a trivial consequence of the fact that $z_{\mathcal{F}}(X)$ is finer than X .

□

To the best of our knowledge, we can't do away with the dependency on \mathcal{F} : one can exhibit a Polish space X with two distinct bases \mathcal{F}, \mathcal{G} such that $z_{\mathcal{F}}(X) \neq z_{\mathcal{G}}(X)$. Despite this apparent lack of canonicity, any Polish topology is entirely determined by its collection of zero-dimensional refinements⁶:

Theorem 3.5 *Any Polish space X has the final topology for the family $\{id_{\mathcal{F}} : z_{\mathcal{F}}(X) \rightarrow X\}_{\mathcal{F}}$ of all the (continuous) identity maps from its zero-dimensionalisations, where \mathcal{F} ranges over all the countable bases of X .*

The proof of this theorem relies on the following lemma.

Lemma 3.6 *Let X be a Polish space and $(x_n)_{n \in \mathbb{N}} \rightarrow x$ a convergent sequence in X . Let \mathcal{F} be a countable base for X . $(x_n)_{n \in \mathbb{N}}$ converges to x in $z_{\mathcal{F}}(X)$ if $x \notin \cup_{O \in \mathcal{F}} \partial O$.*

Proof. Recall that a countable base of $z_{\mathcal{F}}(X)$ is $\mathcal{F} \cup \mathcal{F}|_{\mathcal{F}_\delta^c}$. Assume x is not in the boundary of any element $O \in \mathcal{F}$. Let U be a basic open neighbourhood of x in $z_{\mathcal{F}}(X)$. If $U \in \mathcal{F}$ then it is trivial to exhibit the convergence property by referring

⁶ We mention this fact *en passant* but do not use it in the following developments.

to the topology of X only. If not, we have that $U = O \cap D$, where $D = \bigcap_{i=1}^n X \setminus O_i$, $O_i \in \mathcal{F}$; in other terms $x \in (X \setminus \bigcup_{i=1}^n O_i) \cap O$. Observe that since the O_i are open, $\overline{O_i} = O_i \cup \partial O_i$ – therefore, using the initial assumption, we have $x \in (X \setminus \bigcup_{i=1}^n \overline{O_i}) \cap O$, which is an open set in X . The result follows. \square

Proof. (Theorem 3.5) It suffices to prove that for all topological space Y , a function $f : X \rightarrow Y$ is continuous if and only if $f \circ id_{\mathcal{F}} : z_{\mathcal{F}}(X) \rightarrow Y$ is continuous for all countable base \mathcal{F} . The forward implication is trivial. Assume that for all countable base \mathcal{F} , $f \circ id_{\mathcal{F}} : z_{\mathcal{F}}(X) \rightarrow Y$ is continuous. Consider a converging sequence $(x_n)_{n \in \mathbb{N}} \rightarrow x$ in X . It is sufficient to exhibit *one* space $z_{\mathcal{F}}(X)$ where this sequence also converges. Lemma 3.6 gives as a sufficient criterion that x does not belong to ∂O for any $O \in \mathcal{F}$. Let us build such a base. Consider a dense set \mathcal{D} of X . Let $d : X^2 \rightarrow [0, 1]$ be some metric that completely metrises X . Without loss of generality, assume $x \in \mathcal{D}$. Write $r_n \triangleq d(x, d_n)$ for $d_n \in \mathcal{D} \setminus \{x\}$. For all n , take the family of open balls centred on each d_n with rational radii strictly below r_n , e.g. $r_n/3$. Since $\text{diam}(\overline{B(d_n, r_n/3)}) = \text{diam}(B(d_n, r_n/3))$, $x \notin \partial B(d_n, r)$ for $r < r_n/3$. This family still constitutes a neighbourhood base. The countable union of countable sets is countable, therefore it constitutes a countable base of X . \square

Notice that the topologies of $G(X)$ and $G(z_{\mathcal{F}}(X))$ might be different, and there is in general no continuous map from $G(X)$ to $G(z_{\mathcal{F}}(X))$. It should also be emphasised that the “zero-dimensionalisation” of a Polish space is not an innocent operation: for instance if X is compact and non-zero-dimensional then $z_{\mathcal{F}}(X)$ will *never* be compact! However, we have the following powerful analogue to Thm. 3.5:

Theorem 3.7 *For X Polish, $G(X)$ has the final topology for the family of identity maps $\{G(id_{\mathcal{F}}) : G(z_{\mathcal{F}}(X)) \rightarrow G(X)\}_{\mathcal{F}}$ where \mathcal{F} ranges over countable bases of X .*

Proof. As before, it is sufficient to prove that a map $f : G(X) \rightarrow Y$ is continuous if and only if all precompositions $f \circ G(id_{\mathcal{F}})$ are continuous. If f is continuous then the composites clearly also are. Let us consider the reverse implication and suppose that all composites are continuous. Let $(P_n)_{n \in \mathbb{N}} \rightarrow_{G(X)} P$ be a sequence of probabilities converging weakly to P in $G(X)$. It is sufficient to exhibit one \mathcal{F} s.t. $P_n \rightarrow P$ in $G(z_{\mathcal{F}}(X))$. Let us recall the following theorem ([7], Theorem 2.2): *For any Y Polish, let \mathcal{U} be a subset of $\mathcal{B}(Y)$ such that (i) \mathcal{U} is closed under finite intersections (ii) each open set in X is a finite or countable union of elements in \mathcal{U} . If $P_n(A) \rightarrow P(A)$ for all A in \mathcal{U} , then $P_n \rightarrow_Y P$.* Recall that $\text{Boole}(\mathcal{F})$ is a base of $z_{\mathcal{F}}(X)$. This base trivially verifies condition (i) of the previous theorem. It is therefore sufficient to build a base \mathcal{F} of X such that condition (ii) is verified, i.e. $P_n(A) \rightarrow P(A)$ for all $A \in \text{Boole}(\mathcal{F})$. Observe that the P -continuity sets in X form a Boolean algebra ([27], Lemma 6.4). It then suffices to form a base of X included in the Boolean algebra of continuity sets of X , which is always possible: for any point $x \in X$, there can at most be countably many radii ϵ s.t. the open ball $B(x, \epsilon)$ has a boundary with strictly positive mass. \square

3.2 Zero-dimensional Wallman compactifications

Compactifications are topological operations embedding topological spaces into compact spaces. Common examples are the Alexandrov one-point compactifica-

tion (for locally compact spaces) of the Stone-Čech compactification for Tychonoff spaces. In most settings, this embedding is also required to be dense. By choosing the compactification carefully, one can preserve some relevant properties of the starting space – in our case, Polishness and zero-dimensionality.

A well-behaved class of compactifications (that includes Alexandrov and Stone-Čech as special cases) is that of Wallman compactifications. The general method by which one obtains such a compactification from a given topological space X can be decomposed in two steps:

- (i) one first selects a suitable sublattice of the lattice of open sets of X (a *Wallman base*);
- (ii) then, one topologises (in a standard way) the space of maximal ideals of that particular sublattice.

These compactifications are surveyed in Johnstone [21] and (less abstractly) in Beckenstein et al. [5]. Van Mill [30] provides some facts on Wallman compactifications of separable metric spaces. An extensive topos-theoretic perspective is also given by Caramello [12]. In the remainder of this section, we present this compactification method and apply it to the case of Polish zero-dimensional space, yielding a zero-dimensional compactification that we denote by ω . We then highlight its connection with Prop. 3.1 and study those of its properties that are relevant to our goal.

Spaces of maximal ideals

All the material here is standard from the literature on Stone duality for distributive lattices. See e.g. Johnstone [21] for more details.

Proposition 3.8 *Let X be a set and \mathcal{L} be a (distributive) sublattice of the lattice of subsets of X . The space $\max(\mathcal{L})$ has the set of maximal ideals of \mathcal{L} as points and admits subsets of the form $B(O) = \{I \in \max(\mathcal{L}) \mid O \notin I\}$ as a base. Moreover:*

- (i) $\max(\mathcal{L})$ is $T1$ and compact;
- (ii) if \mathcal{L} is furthermore normal as a lattice, i.e. if for all $O_1, O_2 \in \mathcal{L}$ such that $O_1 \cup O_2 = X$, there exists disjoint O'_1, O'_2 such that $O'_1 \subseteq O_1$, $O'_2 \subseteq O_2$ then $\max(\mathcal{L})$ is Hausdorff.

Proof. It suffices to check that the family $B(O)$ where O ranges in \mathcal{L} is indeed a base (i.e. closed under finite intersections). Maximal ideals are by definition proper. Since \mathcal{L} is distributive, maximal ideals on \mathcal{L} are moreover prime: for all $I \in \max(\mathcal{L})$, $B, B' \in \mathcal{L}$, if $B \cap B' \in I$ then either $B \in I$ or $B' \in I$ ([21], I 2.4). Let $B, B' \in \mathcal{L}$ be given. We show $B(O) \cap B(O') = B(O \cap O')$. Consider $I \in B(O) \cap B(O')$: we have $O \notin I$ and $O' \notin I$, therefore (by primality) $O \cap O' \notin I$, which implies $I \in B(O \cap O')$. Conversely, if $I \in B(O \cap O')$ then $O \cap O' \notin I$. Since ideals are downward closed, we must have $O \notin I$ and $O' \notin I$. For the proof of (i) and (ii), see [21], II resp. 3.5 and 3.6. \square

Wallman bases and compactifications

Wallman compactifications are defined as spaces of maximal ideals over Wallman bases, which are particular lattices that are also bases in the topological sense. Here,

we will follow the definition given in [21]:

Definition 3.9 ([21], IV 2.4) Let X be a topological space and let \mathcal{T}_X be its lattice of open sets. A *Wallman base* is a sublattice of \mathcal{T}_X that is a base for X and which verifies:

For all $U \in \mathcal{T}_X$ and $x \in U$, there exists a $V \in \mathcal{T}_X$ such that $X = U \cup V$ and $x \notin V$.

The following lemma is key in considering a space of maximal ideals over a Wallman base as a compactification (see ([21], IV 2.4) for a proof):

Lemma 3.10 *Let X be a topological space and let \mathcal{L} be a Wallman base for X . $\eta_{\mathcal{L}}(x) = \{O \in \mathcal{L} \mid x \notin O\}$ is a maximal ideal of \mathcal{L} . Moreover, if X is T_0 then $\eta_{\mathcal{L}}$ is an embedding into $\max(\mathcal{L})$.*

We are now in position to define Wallman compactifications:

Definition 3.11 Let X be a T_0 space and \mathcal{L} a Wallman base. We denote $\omega_{\mathcal{L}}(X) = \max(\mathcal{L})$ the *Wallman compactification* of X for \mathcal{L} .

Zero-dimensional compactifications

We will now show that taking the inverse limit of the finite partitions of a Polish zero-dimensional space (as in the proof of Prop. 3.1) corresponds – when applied to a *non-compact* Polish zero-dimensional space – to a Wallman compactification of that space, which exhibits very good properties.

Consider a zero-dimensional Polish space Z . In opposition to the compact case, the Boolean algebra $Clo(Z)$ of clopens of Z is not necessarily countably generated: we therefore consider partitions of Z taken in some countable Boolean sub-algebra $\mathcal{C} \subseteq Clo(X)$ such that \mathcal{C} is a (topological) base for Z . Observe that such a base is always trivially a normal Wallman base. In the following, we call such countable Boolean sub-algebras that generate the topology “Boolean bases”. We define:

$$\mathfrak{C}(X) \triangleq \{\mathcal{C} \mid \mathcal{C} \text{ is a countable Boolean base of } X\}$$

We write $\mathcal{I}_{\mathcal{C}}(Z)$ for the directed partial order of clopen partitions of Z taken in $\mathcal{C} \in \mathfrak{C}(X)$. Since \mathcal{C} is countable, so is $\mathcal{I}_{\mathcal{C}}(Z)$. We recall that the construction of $\mathcal{I}_{\mathcal{C}}(Z)$ is described in the proof of Prop. 3.1.

Proposition 3.12 *For $\mathcal{C} \in \mathfrak{C}(Z)$, let $D_{\mathcal{C}} : \mathcal{I}_{\mathcal{C}}^{op}(Z) \rightarrow \mathbf{Pol}_{fin}$ be the diagram of finite clopen partitions of Z seen as discrete spaces, then $\lim D_{\mathcal{C}}$ is a zero-dimensional compactification of Z homeomorphic to $\omega_{\mathcal{C}}(Z)$.*

Proof. Existence and non-emptiness of $\lim D_{\mathcal{C}}$ stems from surjectivity of the bonding maps and countability of \mathcal{C} . Note that $\lim D_{\mathcal{C}}$ is Polish. Zero-dimensionality is an hereditary property, so it only remains to exhibit an homeomorphism with $\omega_{\mathcal{C}}(Z)$. First, observe that since \mathcal{C} is a Boolean algebra, maximal \mathcal{C} -ideals are in one-to-one correspondence with maximal \mathcal{C} -ultrafilters via the complement map: elements of $\lim D_{\mathcal{C}}$ correspond to \mathcal{C} -filters, they are upward closed and codirected by intersection. They are moreover maximal: for any $U \in \lim D_{\mathcal{C}}$ and all $C \in \mathcal{C}$, either $C \in U$ or $C^c \in U$. A basic clopen in $\lim D_{\mathcal{C}}$ is of the form $\pi_i^{-1}(C)$ where $C \in i \in \mathcal{I}_{\mathcal{C}}$, which

correspond to the ultrafilter $\{U \in \lim D_{\mathcal{C}} \mid C \in U\}$. This in turns, through the negation map, correspond to a basic clopen of $\omega_{\mathcal{C}}(Z)$ (see Prop. 3.8). Every basic clopen of $\omega_{\mathcal{C}}(Z)$ similarly correspond to a basic clopen in $\lim D_{\mathcal{C}}$. Therefore, the spaces are homeomorphic, from which we conclude that $\lim D_{\mathcal{C}}$ is a Polish zero-dimensional compactification. \square

As $\omega_{\mathcal{C}}(Z)$ is always a profinite space, Prop. 3.1 ensures there always exists a cofiltered diagram D in \mathbf{Pol}_{fin} such that $\lim D \cong \omega_{\mathcal{C}}(Z)$. We will switch from one point of view to the other freely. We should insist on the fact that our compactification is *not* the Stone-Ćech compactification, as these are in general not metrisable (except when compactifying an already metrisable compact space, obviously). Take for instance the discrete (hence zero-dimensional) Polish space \mathbb{N} : $\beta\mathbb{N}$ has cardinality $2^{2^{\aleph_0}}$ ([33], Theorem 3.2) while Polish spaces have cardinality at most 2^{\aleph_0} . We would obtain Stone-Ćech if we were to take the Wallman compactification over the full lattice of open sets, however. $\omega_{\mathcal{C}}(Z)$ enjoys a property reminiscent of Stone-Ćech:

Proposition 3.13 *Let Z be a Polish zero-dimensional space. For each continuous map $f : Z \rightarrow K$ to a compact zero-dimensional space K , there exists a Boolean base \mathcal{C} and a continuous map $\omega_{\mathcal{C}}(f) : \omega_{\mathcal{C}}(Z) \rightarrow K$ such that $\omega_{\mathcal{C}}(f) \circ \eta_{\mathcal{C}} = f$, where $\eta_{\mathcal{C}} : Z \rightarrow \omega_{\mathcal{C}}(Z)$ is the embedding of Z into its compactification.*

Proof. Prop. 3.1 entails that there exists a *ccd* $D_K : I^{op} \rightarrow \mathbf{Pol}_{fin}$ s.t. $K \cong \lim D_K$, with limit cone $(\lim D_K, \{\pi_i : \lim D_k \rightarrow D_k(i)\}_{i \in I})$. Note that by continuity of $\pi_i \circ f$, each finite clopen partition of K induces a finite clopen partition of X . By choosing a Boolean base of clopens \mathcal{C} of Z that contains $f^{-1}(Clo(K))$, we can exhibit a compactification $\omega_{\mathcal{C}}(Z)$ with an associated cone $(\omega_{\mathcal{C}}(Z), \{\lambda_i : \omega_{\mathcal{C}}(Z) \rightarrow D_K(i)\})$ and therefore an unique map $\omega_{\mathcal{C}}(f) : \omega_{\mathcal{C}}(Z) \rightarrow K$ such that $\omega_{\mathcal{C}}(f) \circ \eta_{\mathcal{C}} = f$. \square

Corollary 3.14 *For any continuous $f : Z \rightarrow Z'$ between zero-dimensional spaces, there exists Boolean bases $\mathcal{C}, \mathcal{C}'$ of respectively Z and Z' such that there exists a map $\omega_{\mathcal{C}\mathcal{C}'}(f) : \omega_{\mathcal{C}}(Z) \rightarrow \omega_{\mathcal{C}'}(Z')$ verifying $\omega_{\mathcal{C}\mathcal{C}'}(f) \circ \eta_{\mathcal{C}} = \eta_{\mathcal{C}'} \circ f$.*

Zero-dimensional Polish Wallman compactifications were considered in [2], which however does not state Prop. 3.13.

3.3 Projective limit measures on zero-dimensional compactification

For Z Polish zero-dimensional, the developments of Sec. 3.2 allow us to map any measure in $G(Z)$ to $G(\omega_{\mathcal{C}}(Z))$ (for any choice of a Boolean base \mathcal{C}) through $G(\eta_{\mathcal{C}})$. Crucially, thanks to Lemma 2.1 this is a faithful operation.

Therefore any measure on Z can be obtained, up to isomorphism, as a projective limit of finitely supported measures. However, as pointed out before, the converse operation is the difficult one. Let D be a diagram such that $\omega_{\mathcal{C}}(Z) \cong \lim D$ and $\{P_i\}_i \in \lim G \circ D$ a projective family of finitely supported probabilities. There is in general no way to assert that the corresponding projective limit probability $P \in G(\omega_{\mathcal{C}}(Z))$ obtained through the Bochner extension theorem restricts to $G(Z)$. We delineate the conditions under which a probability can be restricted to a subspace and propose a simplification of previous arguments (see [26]), based on the

properties of the Giry monad. Note that the results to follow are not specific to zero-dimensional spaces.

Polish subspaces of Polish spaces are always G_δ sets (and conversely, see [23], 3.11), hence Borel sets. This allows for a simple restriction criterion.

Proposition 3.15 *Let $P \in M(Y)$ be a finite measure on a Polish space Y and let $X \subseteq Y$ be a Polish subspace (hence a G_δ in Y). The restriction of P to X , defined as the set function $P|_X \triangleq (B \in \mathcal{B}(Y) \cap X) \mapsto P(B)$, verifies $P|_X \in G(X)$ if and only if $P(X) = 1$.*

Proof. $P|_X$ is trivially a finite measure on the trace σ -algebra. We observe that $\mathcal{B}(X) = \mathcal{B}(Y) \cap X$: this is a consequence of Theorem 15.1 in [23] (essentially, this follows from the Borel isomorphism theorem for Polish spaces), therefore $P|_X \in M(X)$. Since $P(X) = 1$ and $X \in \mathcal{B}(Y)$, $P|_X(X) = 1$ and $P|_X \in G(X)$. The converse is easy. \square

This criterion lifts to “higher-order” probabilities, that is probabilities over spaces of probabilities, thanks to the multiplication of the Giry monad. The following theorem states that such a higher order probability measure restricts to a subspace if and only if it restricts *in the mean*. This is essentially Theorem 1.1 in [26].

Theorem 3.16 *For all $X \subseteq Y$ Polish spaces and all $P \in G^2(Y)$ we have $P|_{G(X)} \in G^2(X)$ if and only if $(\mu_Y(P))|_X \in G(X)$.*

Proof. The forward implication is trivial. By Lemma 2.1, $G^2(X)$ is a subspace of $G^2(Y)$. By Prop. 3.15, it is sufficient to prove that $P(G(X)) = 1$. By assumption that $\mu(P)|_X \in G(X)$ and Prop. 3.15, we have that $\mu(P)(X) = 1$, which unfolds as $\int_{G(Y)} EV_X dP = 1$. So it suffices to prove that $\int_{G(Y)} EV_X dP = 1 \Rightarrow P(G(X)) = 1$. Assume $P(G(X)) < 1$, then there must exist a Borel set $A \subseteq G(Y) \setminus G(X)$ with $P(A) > 0$. Any probability $p \in A$ assigns positive measure to some Borel set $B \subseteq Y \setminus X$, therefore $\int_A EV_X dP < 1$. \square

4 The Dirichlet process

The Dirichlet process stands out among other Bayesian methods in that the prior and posterior distributions are second order probabilities, that is elements of $G^2(X)$. Learning becomes an operation of type $X \rightarrow G^2(X) \rightarrow G^2(X)$, mapping some evidence in X and a prior in $G^2(X)$ to a posterior in $G^2(X)$, and it can be proved that the second-order stochastic process induced by sampling from identically and independently distributed random variable will converge (in Kullback-Leibler divergence, hence in the weak topology [18]) to a singular distribution over the law of the target.

4.1 The Dirichlet distribution

For a fixed finite discrete space X , Dirichlet is a function $\mathcal{D}_X : M^+(X) \rightarrow G^2(X)$, the parameter in $M^+(X)$ representing the initial prior as well as the degree of certainty about this prior (encoded in its total mass). As we highlight below, \mathcal{D}_X is *continuous*

and verifies other properties, among which *naturality* and *normalisation*. Some of the material on the finitary Dirichlet distribution contained in this section can be found (presented differently) in e.g. [17]. In the following we take $X = \{x_1, \dots, x_n\}$ to be a finite discrete space of cardinality n .

Definition of \mathcal{D}_X

For $Q \equiv (q_1, \dots, q_n) \in M^+(X)$, $\mathcal{D}_X(Q)$ admits a (continuous) density $d_X(Q)$ w.r.t. the $n - 1$ dimensional Lebesgue measure given by

$$d_X(Q)(p_1, \dots, p_{n-1}) \triangleq \frac{\Gamma(\sum_i q_i)}{\prod_i \Gamma(q_i)} \prod_{1 \leq i < n} p_i^{q_i-1} \left(1 - \sum_{1 \leq i < n} p_i\right)^{q_n-1} \quad (1)$$

for all $(p_i)_{i < n} \in \Delta_n \cap (0, 1)^{n-1}$, with $d_X(Q) = 0$ elsewhere. In Equation 1, $\Gamma(q)$ is the continuous gamma function, verifying $\Gamma(n) = (n - 1)!$ for n integer and defined as $\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx$. It is clear that for any bounded continuous function $f \in C_b(\Delta_n)$, $\mathcal{D}_X(Q)(f) = \int_{\Delta_n} f d_X dx^n$ varies continuously with Q , therefore \mathcal{D}_X is continuous in the weak topology.

Γ representation

A useful alternative representation of $\mathcal{D}_X(Q)$ relies on sampling from Γ distributions. For parameters $q, r \in (0, \infty)$, $\Gamma(q, r)$ (not to be mistaken with the Γ function!) is an element of $G(0, \infty)$ with continuous density $f_{q,r}(x) = \frac{x^{q-1} e^{-x/r}}{r^q \Gamma(q)}$. Using the identity $(0, \infty)^n = M^+(X)$ allows to write that for any $Q \equiv (q_1, \dots, q_n) \in M^+(X)$ and $R \equiv (r_1, \dots, r_n) \in M^+(X)$, the product probability $\otimes_{0 \leq i \leq n} \Gamma(q_i, r_i)$ belongs in $G(M^+(X))$. Recall that ν_X is the normalisation map (Sec. 2.1).

Proposition 4.1 (Γ representation)

For all $(q_i)_{1 \leq i \leq n} \in M^+(X)$, $d_X(q_1, \dots, q_n) = \frac{d}{d\lambda}(G(\nu_X)(\otimes_i \Gamma(q_i, 1)))$,

where $\frac{dP}{d\lambda}$ denotes the density of P with respect to Lebesgue.

Proof. We write $x_n \triangleq 1 - \sum_{i < n} x_i$. For any $Q \equiv (q_1, \dots, q_n) \in M^+(X)$, the function ν_X maps the half-line $\{tQ \mid t \in \mathbb{R}_{>0}\}$ to its unique point of intersection with Δ_n . This yields a change of variables $T(t, q_1/|Q|, \dots, q_{n-1}/|Q|) = Q$ with determinant t^{n-1} . Expressing the density of $\otimes_i \Gamma(q_i, 1)$ in these new coordinates and integrating over t to obtain the density at the point of intersection, one obtains:

$$\begin{aligned} \frac{d}{d\lambda}(G(\nu_X)(\otimes_i \Gamma(q_i, 1)))(x_1, \dots, x_{n-1}) &= \int_0^\infty dt \prod_{i \leq n} \frac{(tx_i)^{q_i-1} e^{-tx_i}}{\Gamma(q_i)} t^{n-1} \\ &= \left[\prod_{i \leq n} \frac{x_i^{q_i-1}}{\Gamma(q_i)} \right] \int_0^\infty t^{\sum_i q_i-1} e^{-t} dt \end{aligned}$$

The integral identity $\int_0^\infty x^{b-1} e^{-ax} = a^{-b} \Gamma(b)$ (obtained by integration by parts, or see [20] 3.381.4) yields the desired result. \square

A more detailed proof can be found in [1], Sec. 2.3 or [3], Chap. 27. For an extensive survey of other representations of Dirichlet, see [17].

Extension of \mathcal{D} to non-zero measures

Both the definition in Eq. 1 and the Γ representation have the drawback of being confined to strictly positive measures $M^+(X)$ (since Γ is not defined at 0). M^+ is not a functor unless one restricts to the category of Polish spaces and maps with dense ranges. This is a serious drawback, as strict positivity of a measure is a topological property: in particular, it is not preserved by zero-dimensionalisation of the underlying space! We therefore continuously extend \mathcal{D}_X to the better-behaved nonzero measures, so that it verifies for $Q \in M^+(X \setminus \{x_i\})$:

$$\mathcal{D}_X(M^*(e_i)(Q)) \triangleq G^2(e_i)(\mathcal{D}_{X \setminus \{x_i\}}(Q)) \quad (2)$$

where $e_i : X \setminus \{x_i\} \rightarrow X$ is the inclusion (observe that $M^*(e_i)(Q)(x_i) = 0$). The following proposition asserts that this extension is well-defined as a continuous map.

Proposition 4.2 *Let $Q \in M^+(X \setminus \{x_i\})$ be given. Let δ_{x_i} be the singular probability measure at x_i . Then $\mathcal{D}_X(M^*(e_i)(Q) + \epsilon \delta_{x_i})$ weakly converges to $G^2(e_i)(\mathcal{D}_{X \setminus \{x_i\}}(Q))$ as $\epsilon \rightarrow 0$.*

Proof. We consider w.l.o.g. the case $|X| = 3$, $X = \{x_1, x_2, x_3\}$, $x_i = x_3$. Let $f \in U_b(G(X))$ be given. Still w.l.o.g., take $\|f\| = 1$ and non-negative. In the interest of conciseness, we write $\Gamma_{\otimes}(q_1, q_2, q_3) \triangleq \Gamma(q_1, 1) \otimes \Gamma(q_2, 1) \otimes \Gamma(q_3, 1)$. Using the Γ representation, we have to prove:

$$I(\mathbb{R}_{\geq 0}^{3*}, \epsilon) \equiv \int_{\mathbb{R}_{\geq 0}^{3*}} (f \circ \nu_X) d\Gamma_{\otimes}(q_1, q_2, \epsilon) \xrightarrow{\epsilon \rightarrow 0} \int_{\mathbb{R}_{\geq 0}^{2*}} (f \circ G(e_3) \circ \nu_{x_1, x_2}) d\Gamma_{\otimes}(q_1, q_2) \equiv J(\mathbb{R}_{\geq 0}^{2*})$$

Note that $(f \circ \nu_X)(p_1, p_2, p_3) = f(p_1 / \sum_i p_i, p_2 / \sum_i p_i, p_3 / \sum_i p_i)$ while $(f \circ G(e_3) \circ \nu_{x_1, x_2})(p_1, p_2) = f(p_1 / (p_1 + p_2), p_2 / (p_1 + p_2), 0)$. ν_X is uniformly continuous. Uniform continuity of $f \circ \nu_X$ implies that for all $\epsilon_1 > 0$, there exists an $\eta > 0$ such that for all $(a, b) \in \mathbb{R}_{\geq 0}^{2*}$, for all $c \in [0, \eta]$, $|(f \circ \nu_X)(a, b, c) - (f \circ G(e_3) \circ \nu_{x_1, x_2})(a, b)| < \epsilon_1$. Therefore,

$$\begin{aligned} & I(\mathbb{R}_{\geq 0}^{2*} \times [0, \eta], \epsilon) \\ & \leq \epsilon_1 + \int_{\mathbb{R}_{\geq 0}^{2*} \times [0, \eta]} f(p_1 / \sum_i p_i, p_2 / \sum_i p_i, 0) d[\Gamma_{\otimes}(q_1, q_2, \epsilon)](p_1, p_2, p_3) \\ & \leq \epsilon_1 + \int_{\mathbb{R}_{\geq 0}^{2*}} f(p_1 / \sum_i p_i, p_2 / \sum_i p_i, 0) \left(\int_{[0, \eta]} d\Gamma(\epsilon, 1)(p_3) \right) d[\Gamma_{\otimes}(q_1, q_2)](p_1, p_2) \\ & \leq \epsilon_1 + J(\mathbb{R}_{\geq 0}^{2*}) \end{aligned}$$

One gets the symmetric inequality $J(\mathbb{R}_{\geq 0}^{2*}) - \epsilon_1 \leq I(\mathbb{R}_{\geq 0}^{2*} \times [0, \eta], \epsilon)$ by the exact

same process. As for the other half of the integral,

$$\begin{aligned}
& I(\mathbb{R}_{\geq 0}^{2*} \times (\eta, l], \epsilon) \\
& \leq \int_{\mathbb{R}_{\geq 0}^{2*}} \int_{(\eta, l]} f(p_1 / \sum_i p_i, p_2 / \sum_i p_i, p_3 / \sum_i p_i) \frac{p_3^{1-\epsilon} e^{-p_3}}{\Gamma(\epsilon)} d\Gamma_{\otimes}(q_1, q_2) \\
& \leq \epsilon \int_{\mathbb{R}_{\geq 0}^{2*}} \int_{(\eta, l]} f(p_1 / \sum_i p_i, p_2 / \sum_i p_i, p_3 / \sum_i p_i) \frac{p_3^{1-\epsilon} e^{-p_3}}{\Gamma(1+\epsilon)} dp_3 d\Gamma_{\otimes}(q_1, q_2) \\
& \leq \epsilon K(\eta, \epsilon)
\end{aligned}$$

where we used the identity $\Gamma(1+\epsilon) = \epsilon\Gamma(\epsilon)$ on the third line and $K(\eta, \epsilon)$ is a bounded quantity. Therefore, $I(\mathbb{R}_{\geq 0}^{3*}, \epsilon)$ converges to $J(\mathbb{R}_{\geq 0}^{2*})$ as $\epsilon \rightarrow 0$. This concludes the proof. \square

Naturality of \mathcal{D}

As we are going to show, \mathcal{D} is a natural transformation from M^* to G^2 . This property, usually called *aggregation* in the literature, is an indirect consequence of the closure under convolution of the Γ distribution, a property that we assume without proof:

$$\Gamma(q_1, r) * \Gamma(q_2, r) = \Gamma(q_1 + q_2, r) \quad (3)$$

A corollary of Eq. 3 is that the pushforward of a product of Γ distributions is a product of Γ distributions.

Corollary 4.3 *Consider $f : X \rightarrow Y$ surjective for X, Y finite and discrete, then:*

$$G(M^+(f))(\otimes_{x \in X} \Gamma(q_x, r)) = \otimes_{y \in Y} \Gamma(\sum_{x \in f^{-1}(y)} q_x, r)$$

Proof. It suffices to consider the case of $|X| = 2, |Y| = 1$, in which case $M^+(f) : M^+(X) \rightarrow M^+(Y)$ is simply the addition and $G(M^+(f))$ is by definition the convolution! Eq. 3 then yields the claim. An induction on $|f^{-1}(y)|$ for each $y \in f(X)$ allows to conclude in the general case. \square

Proposition 4.4 $\mathcal{D} \triangleq X \mapsto \mathcal{D}_X$ is a natural transformation $\mathcal{D} : M^* \Rightarrow G^2$ when M^* and G^2 are restricted to \mathbf{Pol}_{fin} .

Proof. It suffices to show that for any $f : X \rightarrow Y$ a (continuous) function between finite discrete spaces, one has $G^2(f) \circ \mathcal{D}_X = \mathcal{D}_Y \circ M^*(f)$. First of all, Prop. 4.2 allows us to restrict our attention to f surjective and $Q = x_i \mapsto q_i \in M^+(X)$ strictly positive. Then:

$$\begin{aligned}
(\mathcal{D}_Y \circ M^+(f))(Q) &= G(\nu_Y)(\otimes_{y \in Y} \Gamma(\sum_{x_i \in f^{-1}(y)} q_i, 1)) && \text{(Prop. 4.1)} \\
&= G(\nu_Y)((G \circ M^+)(f)(\otimes_{x_i \in X} \Gamma(q_i, 1))) && \text{(Cor. 4.3)} \\
&= G(\nu_Y \circ M^+(f))(\otimes_{x_i \in X} \Gamma(q_i, 1)) \\
&= G(G(f) \circ \nu_X)(\otimes_{x_i \in X} \Gamma(q_i, 1)) && \text{(Lemma 2.3)} \\
&= (G^2(f) \circ \mathcal{D}_X)(Q) && \text{(Prop. 4.1)}
\end{aligned}$$

\square

Example 4.5 Let us verify naturality on a simple case: take $X = \{1, 2, 3\}$, $Y = \{1, 2\}$, $f(1) = 1, f(2) = f(3) = 2$ and $Q = i \mapsto q_i$. By definition, $G(f)(p_1, p_2, p_3) = (p_1, p_2 + p_3)$. $G(f)^{-1}(p_1, p_2)$ corresponds to the closed line segment $\{(p_1, \alpha, p_2 - \alpha) \mid \alpha \in [0, p_2]\}$, from which we deduce that

$$\begin{aligned} G^2(\mathcal{D}_X(Q))(B) &= \int_{G(f)^{-1}(B)} d_X(Q)(p_1, p_2) dp_1 dp_2 \\ &= \int_B dp_1 \int_0^{1-p_1} d_X(Q)(p_1, p_2) dp_2 \\ &= \int_B dp_1 \frac{\Gamma(q_1 + q_2 + q_3)}{\Gamma(q_1)\Gamma(q_2)\Gamma(q_3)} p_1^{q_1-1} \int_0^{1-p_1} p_2^{q_2-1} (1-p_1-p_2)^{q_3-1} dp_2 \end{aligned}$$

The last integral can be carried out using the generalised binomial theorem on the rightmost term: the integral identity $\int_0^w x^a (w-x)^b dx = w^{a+b+1} \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(a+b+2)}$ holds ([20], 3.191.1). On the other hand, we have:

$$(\mathcal{D}_Y \circ M^*(f))(Q)(B) = \int_B \frac{\Gamma(q_1 + q_2 + q_3)}{\Gamma(q_1)\Gamma(q_2 + q_3)} p_1^{q_1-1} (1-p_1)^{q_2+q_3-1} dp_1$$

So naturality holds.

Normalisation

The Dirichlet distribution $\mathcal{D}_X : M^+(X) \rightarrow G^2(X)$ obeys a consistency relationship called *normalisation*:

$$\mu_X \circ \mathcal{D}_X = \nu_X \tag{4}$$

Let us verify this identity for a parameter $Q = (q_1, \dots, q_n) \in \mathbb{R}_{>0}^n \subseteq M^+(X)$. Observe that Eq. 4 holds when $|X| = 2$, as $\mathcal{D}_X(q_1, q_2)$ degenerates to a $Beta_X(q_1, q_2)$ distribution which is known to have mean $(\frac{q_1}{q_1+q_2}, \frac{q_2}{q_1+q_2})$ (see eg [3], Sec. 16.5 for a definition of the *Beta* distribution and the proof of this property).

In the case of X an arbitrary finite discrete space, let $f_i : X \rightarrow \{x_i, \bullet\}$ be the lumping function verifying $f_i(x_i) = x_i$, $f_i(x_{j \neq i}) = \bullet$. By naturality of μ and \mathcal{D} :

$$\begin{aligned} (\mu_X \circ \mathcal{D}_X)(Q)(x_i) &= (\mu_{\{x_i, \bullet\}} \circ G^2(f_i) \circ \mathcal{D}_X)(Q)(x_i) \\ &= (\mu_{\{x_i, \bullet\}} \circ \mathcal{D}_{\{x_i, \bullet\}} \circ M^+(f_i))(Q)(x_i) \\ &= (\mu_{\{x_i, \bullet\}} \circ Beta_{\{x_i, \bullet\}})(q_i, \sum_{j \neq i} q_j)(x_i) = \frac{q_i}{\sum_j q_j} \end{aligned}$$

4.2 Extension to zero-dimensional Polish spaces

The finite support case is instructive but lacks generality. We proceed to the extension of finitely supported Dirichlet distributions to Dirichlet processes supported by arbitrary zero-dimensional Polish spaces. Our construction preserves both naturality and continuity – in fact, it can be framed as the extension of the natural transformation \mathcal{D} from \mathbf{Pol}_{fin} to \mathbf{Pol}_z , the full subcategory of zero-dimensional Polish spaces and continuous maps. In what follows, we denote by $F|_{\mathbb{C}} : \mathbb{C} \rightarrow \mathbf{Pol}$ the restriction of the domain of some endofunctor $F : \mathbf{Pol} \rightarrow \mathbf{Pol}$ to a subcategory \mathbb{C} of \mathbf{Pol} . When unambiguous, we drop this notation.

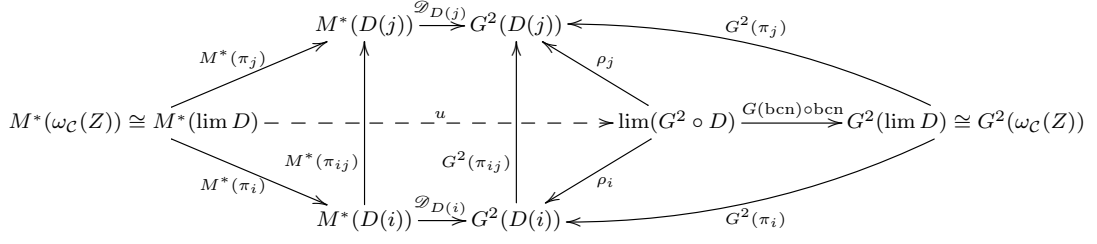


Fig. 1. Construction of $\hat{\mathcal{D}}$ on $\omega_{\mathcal{C}}(Z)$

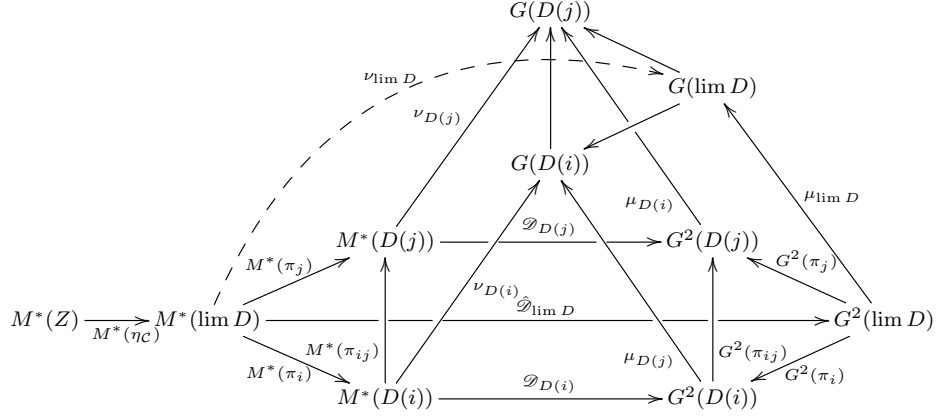


Fig. 2. Commutation of normalisation and Dirichlet averaging

Theorem 4.6 *There exists a unique (up to isomorphism) natural transformation $\hat{\mathcal{D}} : M|_{\mathbf{Pol}_z} \Rightarrow G^2|_{\mathbf{Pol}_z}$ such that $\hat{\mathcal{D}}$ coincides with \mathcal{D} on \mathbf{Pol}_{fin} .*

Proof. We prove existence, naturality and uniqueness.

Existence. For any given choice of a Boolean base \mathcal{C} , let $\eta_{\mathcal{C}} : Z \rightarrow \omega_{\mathcal{C}}(Z)$ be the embedding of a Polish zero-dimensional space Z into its compactification $\omega_{\mathcal{C}}(Z)$ (Lemma 3.10). $\omega_{\mathcal{C}}(Z)$ is compact zero-dimensional so by Prop. 3.1 there exists a *ccd* of finite spaces D such that $\omega_{\mathcal{C}}(Z) \cong \lim D$. Let us construct $\hat{\mathcal{D}}_{\lim D}$, the extension of Dirichlet to $\lim D$ (see Fig. 1). Applying the functor M^* yields a cone $C = (M^*(\lim D), \{M^*(\pi_i) : M^*(\lim D) \rightarrow (M^* \circ D)(i)\}_i)$. Applying the finitary Dirichlet \mathcal{D} on the base of this cone yields a *ccd* in $G^2 \circ D$, of which we take the limit, obtaining a terminal cone $T = (\lim G^2 \circ D, \{\rho_i : \lim G^2 \circ D \rightarrow G^2(D(i))\}_i)$. By naturality of \mathcal{D} , the cone C extends to a cone $C' = (M^*(\lim D), \{\mathcal{D}_{D(i)} \circ M^*(\pi_i) : M^*(\lim D) \rightarrow (G^2 \circ D)(i)\}_i)$. By universality of T , there exists a unique morphism $u : M^*(\lim D) \rightarrow G^2(\lim D)$ mapping C' to T . The Bochner extension theorem (Thm 2.5) yields an isomorphism $G(\text{bcn}) \circ \text{bcn} : \lim G^2 \circ D \rightarrow G^2(\lim D)$ (the fact that $G(\text{bcn})$ is an isomorphism is

a consequence of Lemma. 2.1). This yields a morphism

$$\begin{aligned}\hat{\mathcal{D}}_{\lim D} &: M^*(\lim D) \rightarrow G^2(\lim D) \\ \hat{\mathcal{D}}_{\lim D} &= u \circ G(\text{bcn}) \circ \text{bcn}\end{aligned}$$

that trivially coincides with \mathcal{D} when $\lim D$ happens to be finite. In order to conclude the existence part of the extension, we need to show that $\hat{\mathcal{D}}_{\lim D} \circ M^*(\eta_{\mathcal{C}}) : M^*(Z) \rightarrow G^2(\lim D)$ actually ranges in $G^2(\eta_{\mathcal{C}}(Z)) \subseteq G^2(\lim D)$, after which we can set $\hat{\mathcal{D}}_Z \triangleq \hat{\mathcal{D}}_{\lim D} \circ M^*(\eta_{\mathcal{C}})$. By Theorem 3.16, it suffices to check that for any $Q \in M^*(Z)$,

$$(\mu_{\lim D} \circ \hat{\mathcal{D}}_{\lim D} \circ M^*(\eta_{\mathcal{C}}))(Q) \Big|_{\eta_{\mathcal{C}}(Z)} \in G(\eta_{\mathcal{C}}(Z))$$

which by Prop. 3.15 amounts to checking that this measure attributes full measure to $\eta_{\mathcal{C}}(Z)$. We take advantage of the *normalisation property* (Eq. 4) of \mathcal{D} . Thanks to this property and to the naturality of μ , the diagram in Fig. 2 commutes. The Bochner extension theorem entails the universality of the cone $(G(\lim D), \{G(\pi_i)\}_i)$ at the top of the diagram, therefore commutation of the diagram in Fig. 2 entails the existence of a unique morphism from the cone $(M^*(\lim D), \{\nu_{D(i) \circ M^*(\pi_i)}\}_i)$ to $(G(\lim D), \{G(\pi_i)\}_i)$ (morphism represented as a dashed line in Fig. 2). This morphism is no other than the normalisation $\nu_{\lim D} : M^*(\lim D) \rightarrow G(\lim D)$. Therefore,

$$(\mu_{\lim D} \circ \hat{\mathcal{D}}_{\lim D} \circ M^*(\eta_{\mathcal{C}}))(Q) = (\nu_{\lim D} \circ M^*(\eta_{\mathcal{C}}))(Q)$$

Trivially, $M^*(\eta_{\mathcal{C}})(Q)(Z \setminus \eta_{\mathcal{C}}(Z)) = 0$, therefore $(\nu_{\lim D} \circ M^*(\eta_{\mathcal{C}}))(Q)$ is concentrated on $\eta_{\mathcal{C}}(Z)$. Hence, up to isomorphism, $\hat{\mathcal{D}}_Z$ restricts to a morphism $\hat{\mathcal{D}}_Z : M^*(Z) \rightarrow G^2(Z)$. This concludes the proof of existence.

Naturality. For any map $f : Z \rightarrow Z'$ between zero-dimensional Polish spaces, we must prove $\hat{\mathcal{D}}_{Z'} \circ M^*(f) = G^2(f) \circ \hat{\mathcal{D}}_Z$. By Corollary 3.14, we can reduce the task to the case of a morphism $\omega_{\mathcal{C}'}(f) : \omega_{\mathcal{C}}(Z) \rightarrow \omega_{\mathcal{C}'}(Z')$ between compact zero-dimensional spaces (see Fig. 3a). It remains to prove $\hat{\mathcal{D}}_{\omega_{\mathcal{C}'}(Z')} \circ M^*(\omega_{\mathcal{C}'}(f)) = G^2(\omega_{\mathcal{C}'}(f)) \circ \hat{\mathcal{D}}_{\omega_{\mathcal{C}}(Z)}$. By Prop. 3.1, $\omega_{\mathcal{C}}(Z) \cong \lim D_Z$ and $\omega_{\mathcal{C}'}(Z') \cong \lim D_{Z'}$ where D_Z and $D_{Z'}$ are their respective finite discrete quotient *ccds*. Let us write $(\omega_{\mathcal{C}'}(Z'), \{\pi_i : \omega_{\mathcal{C}'}(Z') \rightarrow D_{Z'}(i)\}_i)$ the terminal cone corresponding to $D_{Z'}$. The universal property of this limit cone allows to reduce the problem to the commutation of the diagram in Fig. 3b:

$$\begin{aligned}\hat{\mathcal{D}}_{Z'} \circ M^*(f) = G^2(f) \circ \hat{\mathcal{D}}_Z &\Leftrightarrow \forall i, G^2(\pi_i) \circ \hat{\mathcal{D}}_{Z'} \circ M^*(f) = G^2(\pi_i) \circ G^2(f) \circ \hat{\mathcal{D}}_Z \\ &\Leftrightarrow \forall i, \mathcal{D}_{D_{Z'}(i)} \circ M^*(\pi_i) \circ M^*(f) = G^2(\pi_i) \circ G^2(f) \circ \hat{\mathcal{D}}_Z\end{aligned}$$

As already argued in the proof of Prop. 3.13, any finite discrete clopen partition of $\omega_{\mathcal{C}'}(Z')$ induces a finite discrete clopen partition of $\omega_{\mathcal{C}}(Z)$ since the two spaces are related by the continuous function $\omega_{\mathcal{C}'}()$, therefore the diagram in Fig. 3b commutes.

Uniqueness. Assume there exists two distinct natural transformations $\hat{\mathcal{D}}, \hat{\mathcal{D}}' : M^*(Z) \rightarrow G^2(Z)$ that coincide with \mathcal{D} on finite spaces. It is clear that it is enough

$$\begin{array}{ccc}
M^*(Z) & \xrightarrow{M^*(f)} & M^*(Z') \\
M^*(\eta_C(Z)) \downarrow & & \downarrow M^*(\eta_{C'}(Z')) \\
M^*(\omega_C(Z)) & \xrightarrow{M^*(\omega_{CC'}(f))} & M^*(\omega_{C'}(Z')) \\
\hat{\mathcal{D}}_{\omega_C(Z)} \downarrow & & \downarrow \hat{\mathcal{D}}_{\omega_{C'}(Z')} \\
G^2(\omega_C(Z)) & \xrightarrow{G^2(\omega_{CC'}(f))} & G^2(\omega_{C'}(Z'))
\end{array}
\quad
\begin{array}{ccc}
M^*(\omega_C(Z)) & \xrightarrow{M^*(\pi_i \circ \omega_{CC'}(f))} & M^*(D_{Z'}(i)) \\
\hat{\mathcal{D}}_{\omega_C(Z)} \downarrow & & \downarrow \mathcal{D}_{D_{Z'}(i)} \\
G^2(\omega_C(Z)) & \xrightarrow{G^2(\pi_i \circ \omega_{CC'}(f))} & G^2(D_{Z'}(i))
\end{array}$$

(a) Reducing naturality to the case of compact zero-dim. Polish spaces.

(b) Finitary case.

to exhibit a contradiction in the case of Z compact zero-dimensional Polish. We refer to Fig. 1 for the notations. Let D be a ccd of finite spaces such that $Z \cong \lim D$, with canonical projections $\pi_i : \lim D \rightarrow D(i)$. By assumption, there must exist a measure $Q \in M^*(Z)$ such that $\hat{\mathcal{D}}(Q) \neq \hat{\mathcal{D}}'(Q)$. But both $\hat{\mathcal{D}}$ and $\hat{\mathcal{D}}'$ verify (by assumption of naturality and consistency with the finitary case) the equalities $G^2(\pi_i) \circ \hat{\mathcal{D}}'_Z = \mathcal{D}_{D(i)} \circ M^*(\pi_i) = G^2(\pi_i) \circ \hat{\mathcal{D}}_Z$ for all i . Therefore, Q induces through $\hat{\mathcal{D}}$ and $\hat{\mathcal{D}}'$ the same projective family of finite-dimensional Dirichlet distributions $\{\mathcal{D}_{D(i)} \circ M^*(\pi_i)(Q)\}_i$, which yields (by unicity of extensions, see Theorem 2.5) a contradiction. \square

4.3 Extension to arbitrary Polish spaces

Let X be an arbitrary Polish space. As shown in Theorem 3.7, $G(X)$ has the final topology for the family of identity maps $\{G(id_{\mathcal{F}}) : G(z_{\mathcal{F}}(X)) \rightarrow G(X)\}_{\mathcal{F}}$ where \mathcal{F} ranges over countable bases of X . In order to harness this theorem, we need the following fact:

Lemma 4.7 *Let X be Polish and $z_{\mathcal{F}}(X), z_{\mathcal{G}}(X)$ be two zero-dimensional refinements as constructed in Prop. 3.2. Then $\mathcal{D}_{z_{\mathcal{F}}(X)}$ and $\mathcal{D}_{z_{\mathcal{G}}(X)}$ are equal in **Set**.*

Proof. The set of countable bases of X is directed by union. Let us write $\mathcal{H} \equiv \mathcal{F} \cup \mathcal{G}$. The (continuous) identity functions $id_{\mathcal{F}\mathcal{H}} : z_{\mathcal{H}}(X) \rightarrow z_{\mathcal{F}}(X)$ and $id_{\mathcal{G}\mathcal{H}} : z_{\mathcal{H}}(X) \rightarrow z_{\mathcal{G}}(X)$ lift to identity functions $G^2(id_{\mathcal{F}\mathcal{H}}), G^2(id_{\mathcal{G}\mathcal{H}})$, and similarly for the functor M^* . Therefore, the commutation relation $G^2(id_{\mathcal{F}\mathcal{H}}) \circ \mathcal{D}_{z_{\mathcal{H}}(X)} = \mathcal{D}_{z_{\mathcal{F}}(X)} \circ M^*(id_{\mathcal{F}\mathcal{H}})$ boils down in **Set** to the equality of $\mathcal{D}_{z_{\mathcal{F}}(X)}$ and $\mathcal{D}_{z_{\mathcal{G}}(X)}$ (and similarly for \mathcal{G}). \square

Finally, we have:

Theorem 4.8 *There exists a unique (up to isomorphism) natural transformation $\hat{\mathcal{D}} : M^* \Rightarrow G^2$ such that $\hat{\mathcal{D}}$ coincides with \mathcal{D} on **Pol**.*

Proof. Let X be a Polish space. Consider the family $\{z_{\mathcal{F}}(X)\}_{\mathcal{F}}$ of its zero-dimensional refinements, as constructed in Prop. 3.2. For each $z_{\mathcal{F}}(X)$, Theorem 4.6 asserts the existence of a continuous Dirichlet map $\hat{\mathcal{D}}_{z_{\mathcal{F}}(X)} : M^*(z_{\mathcal{F}}(X)) \rightarrow G^2(z_{\mathcal{F}}(X))$, which extends by continuity of the identity and functoriality to a continuous map

$$G^2(id_{\mathcal{F}}) \circ \hat{\mathcal{D}}_{z_{\mathcal{F}}(X)} : M^*(z_{\mathcal{F}}(X)) \rightarrow G^2(X) \quad (5)$$

By Lemma 4.7, all these maps coincide in **Set**. Theorem 3.7 allows to conclude. \square

5 Conclusion

Our construction of the Dirichlet process in categorical style subsumes existing ones [17,26] while establishing continuity and naturality. However, further work, which we intend to pursue right away, is required to consolidate our understanding of the finitary approximation framework we have built for higher-order probabilities. The Giry monad can be generalised from **Pol** to the category of Tychonoff spaces, however our construction relies heavily on the properties of Polish spaces: for instance we use the fact that zero-dimensional Polish spaces are Borel sets of their compactifications (Prop. 3.15); the measurable selection theorem used in Lemma 2.1 also requires the spaces considered to be Polish. The process by which we rebuild Dirichlet relies on some simple properties of Γ distributions. Naturality is a consequence of closure of Γ under convolution (a particular case of *infinite divisibility* also exhibited by e.g. normal distributions), and the fact that Dirichlet restricts to X , which is only a subset of its compactification wX , follows from the *normalisation* property (see §4.1). By axiomatising these properties, we can generalise our main result. However, it remains to be seen whether other interesting distributions on $\mathbb{R}_{>0}$ fit the conditions and generate Dirichlet-like processes.

Beyond the immediate questions above, we can return to the less immediate goals expounded on in the introduction, namely higher-order learning using uncertain chains of Dirichlet type. Any uncertain Markov chain τ , meaning a morphism $X \rightarrow G^2(X)$ in **Pol**, can be post-composed with the multiplication of G to obtain the “mean” Markov chain of type $X \rightarrow G(X)$. We will investigate the case where τ takes values in Dirichlet processes -focusing on the tractable “uncertain chains of Dirichlet type”. Such chains can be decomposed as $\alpha : X \rightarrow G(X)$ followed by the Dirichlet natural transformation $\mathcal{D}_X : M^*(X) \rightarrow G^2(X)$. The first component α is τ ’s parameterising chain. As $\mu \circ \mathcal{D}_X \circ \alpha$ is the normalised version of α , α is again up to normalisation the mean chain of the uncertain τ . Our construction ensures that τ is continuous by construction. At this stage, it is already possible given $\tau : X \rightarrow G^2(X)$ to quantify the uncertainty at each point by considering moments of the “Kantorovich” random variable $K_x \triangleq (P \in G(X)) \mapsto d_K((\mu \circ \tau)(x), P)$, where K_X is defined over the probability triple $(G(X), \tau(x))$ and d_K metrises $G(X)$. The next step is to adapt the Bayesian learning scheme which in the discrete case maps the *prior* $\mathcal{D}_X(Q)$ to the *posterior* $\mathcal{D}_X(Q + s)$, for Q in $M^*(X)$ the current parameter, given s a multiset of observed values in X (seen as a counting measure). Via the projective limit construction, learning can be led at the level of behavioural approximants [13] and a subsidiary goal is to understand how the two levels relate. The second goal consists of in extending the probabilistic Kantorovich metric to uncertain chains (of this specific type) and understand its evolution under learning. Until now we assumed that the state of the system is fully observable, but the above questions should be developed as well in a broader context where the state is only partially and noisily so. In this setting, naturality of \mathcal{D} might allow to compare uncertain chains defined on distinct state spaces by embedding them in some universal Polish space – giving a quantitative account of both the differences

in their respective state spaces *and* their dynamics.

Acknowledgement. We warmly thank Laurent Dufloux and Florence Clerc for their valuable insights, and the anonymous reviewers for their thorough and helpful work.

References

- [1] Bela A Frigyik AK and MR Gupta. Introduction to the Dirichlet distribution and related processes. Technical report, Technical Report UWEEETR-2010-0006, Department of Electrical Engineering, University of Washington., 2010.
- [2] F.G. Arenas and M.A. Snchez-Granero. Wallman compactification and zero-dimensionality. *Divulgaciones Matemáticas*, 7(2):151–155, 1999.
- [3] N. Balakrishnan and V.B. Nevzorov. *A Primer on Statistical Distributions*. Wiley, 2004.
- [4] Taras Banach. The topology of spaces of probability measures, i: Functors p_τ and \hat{p} . *Matematychni Studii*, 1995.
- [5] Edward Beckenstein, Lawrence Narici, and Charles Suffel. *Topological algebras*. Mathematic studies. North-Holland, 1977.
- [6] George M. Bergman. Some empty inverse limits. <https://math.berkeley.edu/~gbergman/papers/unpub/>.
- [7] Patrick Billingsley. *Convergence of Probability Measures*. Wiley, 1968.
- [8] Vladimir I Bogachev. *Measure Theory I*. Springer, 2006.
- [9] Vladimir I Bogachev. *Measure Theory II*. Springer, 2006.
- [10] Francis Borceux and George Janelidze. *Galois theories*. Number 72 in Cambridge studies in advanced mathematics. Cambridge University Press, 2001.
- [11] N. Bourbaki. *Elements de mathématique. Topologie Générale*. Springer, 1971.
- [12] O. Caramello. Gelfand spectra and Wallman compactifications. *ArXiv e-prints*, April 2012.
- [13] Philippe Chaput, Vincent Danos, Prakash Panangaden, and Gordon Plotkin. Approximating Markov Processes by averaging. *Journal of the ACM*, 61(1), January 2014. 45 pages.
- [14] Alexandre Donzé and Oded Maler. Robust satisfaction of temporal logic over real-valued signals. In *Formal modeling and analysis of timed systems*, pages 92–106. Springer Berlin Heidelberg, 2010.
- [15] V.V. Fedorchuk. Functors of probability measures in topological categories. *Journal of Mathematical Sciences*, 91(4):3157–3204, 1998.
- [16] Thomas S Ferguson. A Bayesian analysis of some nonparametric problems. *The annals of statistics*, pages 209–230, 1973.
- [17] Subhashis Ghosal. The Dirichlet process, related priors and posterior asymptotics. In N. L. Hjort et al, editor, *Bayesian Nonparametrics*, pages 36–83. Cambridge University Press, 2010.
- [18] Alison L. Gibbs, Francis, and Edward Su. On choosing and bounding probability metrics. *Internat. Statist. Rev.*, pages 419–435, 2002.
- [19] M. Giry. A categorical approach to probability theory. In B. Banaschewski, editor, *Categorical Aspects of Topology and Analysis*, number 915 in Lecture Notes In Mathematics, pages 68–85. Springer-Verlag, 1981.
- [20] I.S. Gradshteyn and D. Ryzhik. *Table of Integrals, Series, and Products*. Elsevier Science, 2000.
- [21] P.T. Johnstone. *Stone Spaces*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1986.
- [22] Olav Kallenberg. *Foundations of Modern Probability*. Springer, 1997.
- [23] Alexander S Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Text in Mathematics*. Springer, 1995.
- [24] Michel Métivier. Limites projectives de mesures. martingales. applications. *Annali di Matematica*, 1963.

- [25] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, 2003.
- [26] Peter Orbanz et al. Projective limit random probabilities on polish spaces. *Electronic Journal of Statistics*, 5:1354–1373, 2011.
- [27] K.R. Parthasarathy. *Probability Measures on Metric Spaces*. AMS Chelsea Publishing Series. Academic Press, 1972.
- [28] Evgeny B Stukalin, Hubert Phillips III, and Anatoly B Kolomeisky. Coupling of two motor proteins: a new motor can move faster. *Physical review letters*, 94(23):238101, 2005.
- [29] Franck van Breugel and James Worrell. A behavioural pseudometric for probabilistic transition systems. *Theoretical Computer Science*, 331(1):115 – 142, 2005. Automata, Languages and Programming.
- [30] Jan Van Mill. *The infinite-dimensional topology of function spaces*, volume 64. Elsevier, 2002.
- [31] Cedric Villani. *Optimal transport, Old and New*. Grundlehren der mathematischen Wissenschaften. Springer, 2006.
- [32] DanielH. Wagner. Survey of measurable selection theorems: An update. In Dietrich Klzow, editor, *Measure Theory Oberwolfach 1979*, volume 794 of *Lecture Notes in Mathematics*, pages 176–219. Springer Berlin Heidelberg, 1980.
- [33] Russell C. Walker. *The Stone-Čech compactification*. Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer, 1974.
- [34] William Waterhouse. An empty inverse limit. *Proceedings of the American Mathematical Society*, 36(2), 1972.
- [35] David Williams. *Probability with Martingales*. Cambridge University Press, 1991.

A Topological and measurable spaces

We recall some basic facts about topological and measurable spaces.

A.1 Topological spaces

Basic definitions

A *topological space* (X, \mathcal{T}_X) is given by a set X and a set \mathcal{T}_X of *open subsets* of X such that $X \in \mathcal{T}_X$, $\emptyset \in \mathcal{T}_X$ and \mathcal{T}_X is closed under arbitrary unions and finite intersections. A set is *closed* if its complement is open. It is *clopen* if it is both closed and open. A set of subsets $\mathcal{S} \subseteq \mathcal{T}_X$ is a *base* of \mathcal{T}_X whenever it is closed under finite intersections and its closure under arbitrary unions yields \mathcal{T}_X . A set of subsets $\mathcal{S} \subseteq \mathcal{T}_X$ is a *pre-base* of \mathcal{T}_X when its closure under finite intersections is a base of \mathcal{T}_X . The *closure* of a subset $A \subseteq X$ is noted \bar{A} , it is the smallest closed set containing A . Conversely, the *interior* of $A \subseteq X$ is noted $\text{int}(A)$, it is the largest open set contained in A . The *boundary* of A is noted ∂A and is defined as $\partial A \triangleq \bar{A} \setminus \text{int}(A)$. A space (X, \mathcal{T}_X) is *separable* if there exists a countable subset $D \subseteq X$ that is *dense* in X , i.e. $\bar{D} = X$. Except where it might lead to ambiguities, we will omit \mathcal{T}_X and write the space simply X . A map $f : X \rightarrow Y$ between two topological spaces is *continuous* if and only if for all $O_Y \in \mathcal{T}_Y$, we have $f^{-1}(O_Y) \in \mathcal{T}_X$. An *homeomorphism* is a bicontinuous bijection. $Y \subseteq X$ is a *subspace* of X if its opens are of the form $O \cap Y$, for $O \in \mathcal{T}_X$. Topological spaces with continuous maps form a category, noted **Top**.

Initial and final topologies

Let $\mathcal{I} = \{f_i : X \rightarrow (X_i, \mathcal{T}_{X_i})\}_i$ be a family of functions f_i from a set X into topological spaces (X_i, \mathcal{T}_{X_i}) . The initial topology induced by \mathcal{I} is the coarsest topology on X making the f_i continuous. It is defined as the topology $\mathcal{T}_{\mathcal{I}}$ generated by the sets $\bigcup_i \{f_i^{-1}(O) \mid O \in \mathcal{T}_{X_i}\}$. The final topology is defined dually, as the finest topology on X making a family of functions $\mathcal{F} = \{f_i : (X_i, \mathcal{T}_{X_i}) \rightarrow X\}_i$ continuous. A subset $O \subseteq X$ is open if and only if $f_i^{-1}(O) \in \mathcal{T}_{X_i}$ for all i . It is straightforward to check that this defines a topology.

Limits and colimits in **Top** are defined by endowing them with resp. the initial and final topologies on the **Set** limits and colimits. In particular, the topological product is defined as the initial topology on the **Set** product w.r.t. the canonical projections.

Metric and metrisable spaces

A *distance* function on a set X is a function $d : X \times X \rightarrow [0, \infty]$ that obeys the following axioms, $\forall x, y, z \in X$: (i) symmetry: $d(x, y) = d(y, x)$, (ii) $d(x, y) \geq 0$, $d(x, y) = 0$ iff $x = y$, (iii) $d(x, y) \leq d(x, z) + d(z, y)$. Any distance d on X induces the topology of a *metric space*, with base the open balls $B(x, \epsilon) = \{y \mid d(x, y) < \epsilon\}$, for all $x \in X$, $\epsilon > 0$. A metric space is noted (X, d) . A sequence of points $(x_n \in X)_{n \in \mathbb{N}}$ *converges* to a point $x \in X$ if for all $\epsilon > 0$, there exists a $N \in \mathbb{N}$ such that for all $n \geq N$, $d(x_n, x) < \epsilon$. A sequence is *Cauchy* if for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $m, n \geq N$, $d(x_m, x_n) < \epsilon$. A metric space is *complete* if all Cauchy sequences converge. A space is *metrisable* if its topology is generated

by some distance. A space is *completely metrisable* if its topology is generated by some distance that makes it *complete*. A *Polish space* is a separable, completely metrisable space. Polish spaces form a full subcategory of **Top**, noted **Pol**. **Pol** has all countable limits and all countable disjoint unions [11]. It includes the category of finite, discrete topological spaces \mathbf{Pol}_{fin} as a full subcategory (a space X is discrete if $\mathcal{T}_X = \wp(X)$).

Separation conditions

A topological space X is *Hausdorff* if for any two points $x, y \in X$ there exists disjoint open sets O_x, O_y such that $x \in O_x, y \in O_y$. In a Hausdorff space, all singletons are closed. X is *completely regular* if for any closed set $F \subseteq X$ and any point $x \in X \setminus F$, there exists a continuous function $f : X \rightarrow \mathbb{R}$ such that $f(x) = 0$ and $f(y) = 1$ for all $y \in F$. A space is *Tychonoff* if it is completely regular and Hausdorff. All metrisable spaces are Tychonoff.

Compactness

An *open cover* of a space X is a family $\{O_i \in \mathcal{T}_X\}_i$ of open subsets such that $\cup_i O_i = X$. A topological space is *compact* if any open cover of the space has a finite sub-cover. A subset of X is compact if it verifies this property. All finite subsets are compact. The continuous image of a compact set is compact. All spaces we consider will be Hausdorff, accordingly all compact spaces will be implicitly Hausdorff. All compact subspaces of Hausdorff spaces are closed. *Tychonoff's* theorem asserts that an arbitrary product of compact spaces is compact. A continuous bijection between compact (Hausdorff!) spaces is always a homeomorphism.

Zero-dimensional spaces

A topological space is *zero-dimensional* if it has a base of clopen sets. The set of clopen sets of a space X is noted $\mathcal{C}(X)$. It is a Boolean algebra. One easily deduces that zero-dimensional Polish spaces have a countable base of clopen sets. Zero-dimensionality is a hereditary property and is preserved by subspaces.

Compactifications

A *compactification* of a (Tychonoff) topological space X is a compact space Y into which X embeds homeomorphically and such that the closure of X in Y is Y itself (a non-Tychonoff spaces need not embed in its compactification).

A.2 Measurable spaces

A *measurable space* (X, Σ_X) is a set X along a σ -algebra Σ_X , that is a set of subsets of X closed under complements and countable unions that contain X . If X is a topological space we note $\mathcal{B}(X)$ the *Borel* σ -algebra generated from its topology. A map $f : (X, \Sigma_X) \rightarrow (Y, \Sigma_Y)$ between measurable spaces is *measurable* if $f^{-1}(A) \in \Sigma_X$ for all $A \in \Sigma_Y$. If $f : X \rightarrow Y$ is a continuous map, f is also measurable between the corresponding Borel measure spaces. Borel measure spaces arising from Polish spaces verify the “Isomorphism theorem” [23]:

Theorem A.1 *For all X and Y Polish spaces, $\mathcal{B}(X) \cong \mathcal{B}(Y)$ if and only if X and Y have the same cardinality.*

B Proof of Theorem 2.5

This proof is adapted from Metivier [24].

Let $D : I^{op} \rightarrow \mathbf{Pol}$ be our *ccd*, with canonical projections $\pi_i : \lim D \rightarrow D(i)$. Let $\{P_i\}_i \in \lim G \circ D$ be given. We proceed to continuously extend this family to an element $P \in G(\lim D)$. Consider $\mathcal{A} = \cup_{i \in I} \{\pi_i^{-1}(B) \mid B \in \mathcal{B}(D(i))\}$. By directedness, \mathcal{A} is an algebra of $\lim D$ -Borel sets. We define the set function $P_0 : \mathcal{A} \rightarrow [0, 1]$ by $P_0(\pi_i^{-1}(B)) = P_i(B)$. Codirectedness of the family $\{P_i\}_i$ ensures that (i) P_0 is consistent as a function and that (ii) P_0 is finitely additive, therefore P_0 is a *charge*. As P_0 is finite, hence σ -finite, it is sufficient to exhibit that P_0 is σ -additive on \mathcal{A} and the Carathéodory extension theorem ([35], Theorem 1.7) will yield the sought unique projective limit Borel measure. σ -additivity is equivalent to the implication $\forall n, P_0(A_n) \geq \delta \Rightarrow \cap_n A_n \neq \emptyset$ for all $\delta > 0$ and all decreasing sequence of Borel sets $(A_n)_{n \in \mathbb{N}}$ ([8], Prop. 1.3.3).

Let $(A_n)_n$ be such a sequence. Each A_n is by construction of the form $A_n = \pi_i^{-1}(B_{i,n}^*)$ for some $i \in I$, where $B_{i,n}^* \in \mathcal{B}(D(i))$. We map this sequence $(A_n)_n$ to a family $\{B_{c_n} \in \mathcal{B}(D(c_n))\}_{c_n}$ of Borel sets indexed by an increasing sequence $(c_n)_{n \in \mathbb{N}}$, cofinal in I , such that for all n , $A_n = \pi_{c_n}^{-1}(B_{c_n})$ and $B_{c_{n+1}} \subseteq \pi_{c_n c_{n+1}}^{-1}(B_{c_n})$. The cofinal increasing sequence $(c_n)_{n \in \mathbb{N}}$ is constructed by induction on any fixed enumeration of I . By construction, there is some $i_n \in I$ for which $A_n = \pi_{i_n}^{-1}(B_{i_n})$. By cofinality, there exists $c_n \geq i_n$ and by measurability, $B_{c_n} \triangleq \pi_{i_n c_n}^{-1}(B_{i_n})$ is measurable. By directedness, $A_n = \pi_{c_n}^{-1}(B_{c_n})$. Now consider $m \leq n$ with $A_n \subseteq A_m$. We have $A_n = A_n \cap A_m = \pi_{c_n}^{-1}(B_{c_n}) \cap \pi_{c_m}^{-1}(B_{c_m})$. By directedness, $\pi_{c_n}^{-1} = \pi_{c_m c_n}^{-1} \circ \pi_{c_n}^{-1}$ therefore $A_n = \pi_{c_n}^{-1}(B_{c_n} \cap \pi_{c_m c_n}^{-1}(B_{c_m}))$. For n fixed, this generalises to $A_n = \pi_{c_n}^{-1}(B_{c_n}) = \pi_{c_n}^{-1}(\cap_{m \leq n} \pi_{c_m c_n}^{-1}(B_{c_m}))$. Therefore, $B_{c_{n+1}} = \cap_{m \leq n+1} \pi_{c_m c_{n+1}}^{-1}(B_{c_m}) = \cap_{m \leq n+1} (\pi_{c_n c_{n+1}}^{-1} \circ \pi_{c_m c_n}^{-1})(B_{c_m}) \subseteq \pi_{c_n c_{n+1}}^{-1}(\cap_{m \leq n} \pi_{c_m c_n}^{-1}(B_{c_m})) = \pi_{c_n c_{n+1}}^{-1}(B_{c_n})$.

We construct a nonempty (compact!) set K s.t. $K \subseteq A_n$ for all n . By cofinality of $(c_n)_{n \in \mathbb{N}}$, it is sufficient to construct of a family of non-empty compact sets $\{K_{c_n} \subseteq B_{c_n}\}_{n \in \mathbb{N}}$ that is projective, i.e. verifying $\pi_{c_n c_{n+1}}(K_{c_{n+1}}) = K_{c_n}$ for all n . Such a projective family of compact sets can in turn be obtained from a sequence of non-empty compact sets $(K'_{c_n})_{n \in \mathbb{N}}$ verifying $K'_{c_{n+1}} \subseteq \pi_{c_n c_{n+1}}^{-1}(K'_{c_n})$. Indeed, setting, for all m , $K_{c_m} = \cap_{n \geq m} \pi_{c_m c_n}(K'_{c_n})$, we trivially have that K_{c_m} is compact. As an intersection of a decreasing sequence of non-empty compact sets (in a metrisable space), K_{c_m} is also non-empty (this is Cantor's intersection theorem). Moreover, $K_{c_m} = \cap_{n \geq m} \pi_{c_m c_n}(K'_{c_n}) \supseteq \cap_{n \geq m+1} (\pi_{c_m c_{m+1}} \circ \pi_{c_{m+1} c_n})(K'_{c_n}) \supseteq \pi_{c_m c_{m+1}}(\cap_{n \geq m+1} \pi_{c_{m+1} c_n}(K'_{c_n})) = \pi_{c_m c_{m+1}}(K_{c_{m+1}})$. To prove the reverse inclusion, it suffices to show that for all $x \in K_{c_m}$, $\pi_{c_m c_{m+1}}^{-1}(x) \cap K_{c_{m+1}} \neq \emptyset$. We have $\pi_{c_m c_{m+1}}^{-1}(x) \cap K_{c_{m+1}} = \pi_{c_m c_{m+1}}^{-1}(x) \cap (\cap_{n \geq m+1} \pi_{c_{m+1} c_n}(K'_{c_n})) = \cap_{n \geq m+1} (\pi_{c_m c_{m+1}}^{-1}(x) \cap \pi_{c_{m+1} c_n}(K'_{c_n}))$. Notice that $\pi_{c_m c_{m+1}}^{-1}(x) \cap \pi_{c_{m+1} c_n}(K'_{c_n})$ is compact for all n . Since by definition $\pi_{c_m c_{m+1}}^{-1}(x) \subseteq K_{c_{m+1}} \subseteq \pi_{c_{m+1} c_n}(K'_{c_n})$ for all n , this intersection is non-empty.

We have reduced the goal to providing a sequence of non-empty compact sets $(K'_{c_n} \subseteq B_{c_n})_{n \in \mathbb{N}}$ verifying $K'_{c_{n+1}} \subseteq \pi_{c_n c_{n+1}}^{-1}(K'_{c_n})$. Recall that $P(A_n) \geq \delta > 0$ for

all n , which implies $P_{c_n}(B_{c_n}) \geq \delta > 0$ for all n . Finite Borel measures on Polish spaces are *Radon*: for all $P \in G(X)$ with X Polish, for all $B \in \mathcal{B}(X)$, $P(B) = \sup \{P(K) \mid K \subseteq B, K \text{ compact}\}$ ([7], Theorem 1.4); therefore each $P_{c_n} \in G(D(c_n))$ is Radon. We build by induction a sequence $(K'_{c_n})_{n \in \mathbb{N}}$ such that $P_{c_n}(B_{c_n} \setminus K'_{c_n}) < \sum_{k=0}^n \frac{\epsilon}{2^{k+1}}$ and $K'_{c_{n+1}} \subseteq \pi_{c_n c_{n+1}}^{-1}(K'_{c_n})$. For $n = 0$, we obtain K'_{c_0} verifying $P_{c_0}(B_{c_0} \setminus K'_{c_0}) < \epsilon/2$ by application of the Radon property. Our inductive hypothesis consists in the existence of a sequence $(K'_{c_k})_{0 \leq k \leq n}$ having the aforementioned properties. By assumption, $B_{c_{n+1}} \subseteq \pi_{c_n c_{n+1}}^{-1}(B_{c_n})$. We have $\pi_{c_n c_{n+1}}^{-1}(K'_{c_n}) \cap B_{c_{n+1}} = B_{c_{n+1}} \cap [\pi_{c_n c_{n+1}}^{-1}(B_{c_n}) \setminus \pi_{c_n c_{n+1}}^{-1}(K'_{c_n})]$, therefore $P_{c_{n+1}}(B_{c_{n+1}} \setminus \pi_{c_n c_{n+1}}^{-1}(K'_{c_n})) < \sum_{k=0}^n \epsilon/2^{k+1}$. To conclude it suffices to pick, using the Radon property, $K'_{c_{n+1}}$ s.t. $P_{c_{n+1}}((B_{c_{n+1}} \cap \pi_{c_n c_{n+1}}^{-1}(K'_{c_n})) \setminus K'_{c_{n+1}}) < \epsilon/2^{n+2}$. One then has a sequence verifying all the required properties – in particular, its elements are non-empty (since they have positive measure) and they verify $K'_{c_{n+1}} \subseteq \pi_{c_n c_{n+1}}^{-1}(K'_{c_n})$. This concludes the existence and unicity of the measure associated to $\{P_i\}_i \in \lim G \circ D$.

We now prove that this extension is a homeomorphism. Observe that the maps $G(\pi_i) : G(\lim D) \rightarrow G(D(i))$ define a cone over $G \circ D$, therefore there exists an universal (continuous!) mediating map $\text{bcn}^{-1} : G(\lim D) \rightarrow \lim G \circ D$, which associates to $P \in G(\lim D)$ a projective system $\{G(\pi_i)(P)\}_{i \in I}$ of probabilities. As Borel measures are entirely specified by their values on open sets (Lemma 7.1.2, [9]), bcn^{-1} is injective. The uniqueness of the procedure described above ensures that P is precisely the extension corresponding to $\{G(\pi_i)(P)\}_{i \in I}$, therefore bcn^{-1} is surjective. Let us prove continuity of bcn . Consider $(t_n \in \lim G \circ D)_{n \in \mathbb{N}}$ a sequence converging to $t \in \lim G \circ D$, i.e. for all $i \in I$ that $t_n(i)$ weakly converges (in the sense of Sec. 2.1) to $t(i)$, which is equivalent by the “Portmanteau” theorem to strong convergence on $t(i)$ -continuity sets for all i . Let us write $(P_n = \text{bcn}(t_n))_n, P = \text{bcn}(t)$ the projective limit measures of resp. $(t_n)_n, t$. We must prove $P_n \rightharpoonup P$. It can be easily verified, by commutation of the interior and closure operations with topological products, that for all i , if $B \in \mathcal{B}(G(D(i)))$ is a $t(i)$ -continuity set then $\pi_i^{-1}(B)$ is a P -continuity set. Let d_i be a distance compatible with $D(i)$, consider for $x \in D(i)$ the neighbourhood $N_{i,x}(\epsilon) = \{y \in D(i) \mid d_i(x, y) \leq \epsilon\}$. For distinct ϵ_k , $\partial N_{i,x}(\epsilon_k)$ are disjoint. Therefore there cannot be more than a countable family of $\{\epsilon_k > 0\}_k$ such that $t(i)(\partial N_{i,x}(\epsilon_k)) > 0$. We deduce that each prebase open $\pi_i^{-1}(N_{i,x}(\epsilon))$ contains a continuity set. Since continuity sets form an algebra, Corollary 1 to Theorem 2.2 of [7] applies and we conclude that $P_n \rightharpoonup P$. Therefore, bcn is a homeomorphism.